

VOLUME 23 NUMBER 3 SEPTEMBER 1960

SOCIOMETRY

**A JOURNAL OF RESEARCH
IN SOCIAL PSYCHOLOGY**

Articles by

STRYKER and PSATHAS
KELLEY and ARROWOOD
BLAU
HAYS
HOFFMAN and SMITH
MAHAR
TURNER

Published by

THE AMERICAN SOCIOLOGICAL ASSOCIATION

SOCIOMETRY

EDITOR

JOHN A. CLAUSEN

ASSOCIATE EDITORS

EDGAR F. BORGATTA
ORVILLE G. BRIM, JR.
LAUNOR F. CARTER
NATHANIEL L. GAGE
ERVING GOFFMAN

SIDNEY SIEGEL
SHIRLEY A. STAR
HERBERT A. SIMON
FRED L. STRODTBECK
GUY E. SWANSON

EDITORIAL CONSULTANTS

WENDELL BELL
PETER M. BLAU
ROGER W. BROWN
JAMES S. COLEMAN
LEONARD S. COTTRELL, JR.
SANFORD M. DORNBUSCH
FRED E. FIEDLER
ALLAN R. HOLMBERG
ALEX INKELES
HERBERT C. KELMAN
MELVIN L. KOHN
WILLIAM W. LAMBERT

DUNCAN MACRAE, JR.
JAMES G. MARCH
FREDERIC MOSTELLER
ROBERT RADLOW
THEODORE R. SARBIN
STANLEY SCHACHTER
WILLIAM A. SCOTT
JOHN THIBAUT
RALPH H. TURNER
EUGENE A. WEINSTEIN
ROBERT F. WINCH
MARION RADKE YARROW

EXECUTIVE OFFICER

ROBERT BIERSTEDT

Published four times a year

March, June, September, December

by

THE AMERICAN SOCIOLOGICAL ASSOCIATION

at The Boyd Printing Co., Inc., 49 Sheridan Ave., Albany 10, N. Y.

Sociometry was founded in 1937 by J. L. Moreno

Communications for the Editors and all manuscripts should be addressed to John A. Clausen, Institute of Human Development, Institute Annex, 2251 College Avenue, Berkeley 4, California. Business communications should be addressed to Robert Bierstedt, The American Sociological Association, New York University, Washington Square, New York 3, New York.

Copyright © 1960 by THE AMERICAN SOCIOLOGICAL ASSOCIATION

Second class postage paid at Albany, N. Y.

SOCIOMETRY

A JOURNAL OF RESEARCH IN SOCIAL PSYCHOLOGY

VOLUME 23 · NUMBER 3 · SEPTEMBER 1960

CONTENTS

- Research on Coalitions in the Triad: Findings, Problems and Strategy
Sheldon Stryker and George Psathas 217
- Coalitions in the Triad: Critique and Experiment
Harold H. Kelley and A. J. Arrowood 231
- Patterns of Deviation in Work Groups.....*Peter M. Blau* 245
- Psychological Dimensionality and the Distribution of Rank Order Agreement Among Judges.....*William L. Hays* 262
- Some Factors Affecting the Behaviors of Members of Problem-Solving Groups.....*L. Richard Hoffman and Clagett G. Smith* 273
- A Ritual Pollution Scale for Ranking Hindu Castes
Pauline Moller Mahar 292
- Preoccupation with Competitiveness and Social Acceptance Among American and English College Students.....*Ralph H. Turner* 307

SOCIOMETRY

EDITORIAL POLICY

Sociometry is concerned with the entire range of interests and problems represented by research in social psychology. It is the policy of the editors to seek those manuscripts for publication which represent the significant research interests of investigators who are concerned with giving the field of social psychology theoretical structure and reporting research which is clearly focused, well designed, and competently conducted.

While social psychology is presently regarded by most as a field with indeterminate boundaries, it has as its central focus the investigation of the processes and products of social interaction at the interpersonal, intrapersonal, intergroup, and intragroup levels and the development of significant generalizations therefrom. In keeping with the more general meaning of the name of the journal emphasis will be placed on measurement of social behavior. However, this emphasis does not exclude the acceptability of good articles which must rely upon qualitative materials and analyses.

The editors and editorial consultants can be expected to subject manuscripts to rigorous criticism and screening according to the best standards of scientific research and at the same time avoid a sterile orthodoxy which would stultify the communication of creative ideas at the growing edge of the science. Thus the journal will strive to be flexible in its response to the publication needs of its contributors.

It is the intention of the editors to avoid any tendency toward professional provincialism and to invite contributions from any sector of the scientific community which promise to further the objectives of the journal.

Annual subscription rate \$9.00; special rate to members of THE AMERICAN SOCIOLOGICAL ASSOCIATION \$4.50. Single issues \$2.50.

Specifications for Manuscript Preparation for Sociometry may be obtained from the Editorial Office.

Back Issues dated 1937 through 1955 should be ordered from the former publisher, Beacon House, Beacon, New York.

Notice of *change of subscriber's address* should be sent to THE ASSOCIATION four weeks in advance, indicating the old as well as the new address.

Research on Coalitions in the Triad: Findings, Problems and Strategy¹

SHELDON STRYKER AND GEORGE PSATHAS
Indiana University

INTRODUCTION

The research reported here stems from three sources: Theodore Caplow's effort to state a formal theory of coalitions in the triad (2, 1); Vinacke and Arkoff's attempt to test this theory experimentally (5); and Simmel's discussion of *tertius gaudens* (3). Its purposes are, first, to present data which appear to be relevant to hypotheses drawn from Caplow's theory and from Simmel's theorizing; second, to raise questions about the adequacy of the experimental procedure and the data obtained to test certain of these hypotheses.

Caplow examines the model of the triad whose members differ in power, calling "attention to a neglected feature of this model, namely, that the formation of given coalitions depends on the initial distribution of power, and other things being equal, may be predicted to some extent when the initial distribution of power is known." Four assumptions underlie his predictions: "1. Members of a triad may differ in strength. A stronger member can control a weaker member, and will seek to do so. 2. Each member of the triad seeks control over the others. Control over two others is preferred to control over one. Control over one other is preferred to control over none. 3. Strength is additive. The strength of a coalition is equal to the sum of the strengths of its two members. 4. The formation of coalitions takes place in an existing triadic situation, so that there is a precoalition condition in every triad. Any attempt by a stronger member to coerce a weaker member into joining a non-advantageous coalition will provoke the formation of an advantageous coalition to oppose the coercion." ²

Caplow considers six types of triads, differing in initial (pre-coalition) distribution of power. These appear in Table 1 along with his predictions. For example, in triad type 3, he predicts that the coalition formed will include the "weak" man along with (either) one of the two equally "strong." This predic-

¹ Presented to the Social Psychology Section, Ohio Valley Sociological Society, April, 1959. This research was supported by a grant from the U. S. Public Health Service, M-2821(A), Small Grants Program. We are indebted to William Chambliss, research assistant, for his important contributions to this study.

² Caplow has recently extended his theory to apply to diverse "interactive situations," doing so by altering assumption 2. See (1).

TABLE 1
Triad Types and Predicted Coalitions in the Triad

Triad Type	Weights Assigned by Vinacke-Arkoff	Coalitions Predicted by Caplow	Coalitions Predicted by Game Theory
1. $A = B = C$	1-1-1	any	any
2. $A > B, B = C,$ $A < (B + C)$	3-2-2	BC	any
3. $A < B, B = C$	1-2-2	AB or AC	any
4. $A > (B + C)$ $B = C$	3-1-1	none	none
5. $A > B > C,$ $A < (B + C)$	4-3-2	AB or AC	any
6. $A > B > C$ $A > (B + C)$	4-2-1	none	none

tion is based on the observations that neither B nor C can improve upon their pre-coalition relationship of equality, should they join; that the only way in which A can secure an advantage over B or C is through a coalition with one of these; and that if B or C joins A, he retains his superiority over A in the coalition and he improves upon his pre-coalition relationship with the remaining member of the triad.

Vinacke and Arkoff point out that from a strictly "rational" view which they suggest derives from game theory, in which players are presumed to act solely in terms of best strategy to maximize winnings, any coalition is equally likely in triad types 2, 3 and 5. Further, such spoils as are won by a coalition should be divided equally between the pair, since one is as necessary to victory as the other regardless of initial strengths. This is true because, as a matter of fact, anyone failing to enter a coalition cannot win against any combination of the two others. Any player must presume that, should he attempt to remain aloof, the two others will form a coalition to defeat him. To forestall isolation, then, any player threatened with omission from a coalition should be willing to grant another at least half the spoils in order to share in these. Assuming an unlimited bargaining process, there is thus no basis for predicting the greater likelihood of any particular pair coalescing, and no rational reason for any player to settle for less than 50 per cent of the spoils.

To test these contrasting predictions, Vinacke and Arkoff assign appropriate weights, using "weight" as the operational equivalent of "strength" or "power," to three persons, place the trio in a game situation in which weights alone determine outcome, permit coalitions to form during the game, and require an agreement by the coalescing partners on a division of the prize. Illustrative weights are indicated in Table 1. The game uses a modified

parchesi board and a single die. The person, or coalition, winning the game, the object of which is to reach "home" first, receives a reward of 100 points. All players start from the same point, and all move with each throw of the die. Thus, if A is assigned weight 1, B and C weights 2, and the die comes up 5, then A moves 5 spaces, B and C 10 spaces each. Coalescing players move at a rate equal to the sum of their individual weights times the value of the die. If A and B, in this example, coalesce, they move at a rate of three times the value of the die.

Vinacke and Arkoff report the results of 90 trials for each of the triad types. They interpret their findings as supporting what we shall call "perception theory" in contrast to game theory, in that coalition formation apparently is related to the initial strengths of the players at the beginning of the games.³

The present research uses the Vinacke-Arkoff game, focusing on the Type 3 triad, $A < B$, $B = C$. It asks whether new data will replicate the Vinacke-Arkoff findings and so, presumably, strengthen the grounds for the acceptance of a perception theory of coalition formation.⁴ Beyond this, however, there are other questions of interest.

In his discussion of *tertius gaudens*, Simmel contends that the weak man (*tertius*) in a triad including two equally strong others can profit far out of proportion to his intrinsic power by aligning himself with one of the two more powerful members. Further, "the quantity of his power is determined exclusively by the strength which each of (the other members) has relative to the other . . . the only important thing is that his superadded power gives one of them superiority" (3, p. 157). The context in which the weak may acquire maximal profit is, according to Simmel, one of contention between two more powerful persons. If these two are not in contention, the advantage of *tertius* is limited, since he no longer holds the balance of power. Moreover, *tertius's* possibility of deriving maximal profit depends on his freedom to choose which way he will throw his strength. If he is constrained in such choice, his hand is weakened (3, pp. 157-159).

We do not know, from the Vinacke-Arkoff study, whether variations in initial strength of the weak man in a triad relative to the two stronger others

³ For example, in the Type 3 triad, Vinacke and Arkoff report that 24 AB coalitions, 40 AC coalitions, and 15 BC coalitions were formed, with no coalitions forming in 11 games. This distribution departs significantly from a chance distribution, as measured by Chi Square. Further, prize divisions ranging from 1/99 to 29/71 were agreed to in 23 per cent of the coalitions, in 38 per cent the prize division ranged from 30/70 to 49/51, while in 39 per cent of the coalitions the prize division was 50/50. See also the later paper by Vinacke (4), containing additional data for male subjects, and comparing the behavior of females and males in their game situation.

⁴ As will be indicated in the discussion, it now seems clear that, even were the Vinacke-Arkoff findings replicated in detail, there would be, in fact, no grounds for the acceptance of a perception theory in contrast to a game theory of coalition formation.

will make a difference in coalition outcomes. An extension of the Vinacke-Arkoff interpretation of Caplow would say yes. Simmel would say no.

HYPOTHESES

Dealing specifically with the Type 3 triad, we propose to test the following hypotheses:

1. Weak-strong coalitions will occur more frequently than can be expected by chance (from Caplow, Vinacke-Arkoff).
2. In weak-strong coalitions, the division of the prize obtained by the weak man will be less than 50 per cent (from Caplow, Vinacke-Arkoff).
3. Increases in weight of the weak man relative to the two stronger men will result in (a) more coalitions including the weak man, (b) larger proportions of the prize being obtained by the weak man (from Caplow, Vinacke-Arkoff).
4. Contention between the two strong members of the triad will result in larger proportions of the prize being obtained by the weak man (from Simmel).
5. Contention between the weak man and one of the strong men will result in smaller proportions of the prize being obtained by the weak man (from Simmel).

RESEARCH DESIGN

The hypotheses call for a design involving as independent variables weights, contention between the two strong players (which we will call S-S contention), and contention between the weak and one of the strong players (which we will call W-S contention). The dependent variables are frequencies of coalitions and proportions of the prize.

Power weight was assigned randomly to members of the triad, with the weight of the weak man varying among 1, 3, and 5, and the weights of the two strong men always being 6. The weak man was decided by lot at the beginning of each game, the particular person drawing the "weak" weight always designated by the letter "A." To distinguish between the two stronger players, markers of different colors were used.

Contention was created by verbally instructing appropriate players that they "have been enemies of long standing, and cannot form a coalition."

Our instructions "forced" coalitions by having subjects replay tie games, i.e., two persons reaching "home" simultaneously without having formed a coalition required a replay of the game. Vinacke-Arkoff allowed games to be played with no coalitions formed.

The experimental design used was a 12 x 12 Latin Square. Twelve triads

were assigned randomly, within requirements of the Latin Square, to play 12 games, each game characterized by one of two S-S contention conditions (contention absent or present), one of two W-S contention conditions (contention absent or present), and one of three weight conditions.⁵ Subjects were male undergraduates, ages 18-22, recruited in blocks of 36 and assigned randomly to triads.

FINDINGS

Before presenting the Latin Square analysis, hypothesis 1 (weak-strong coalitions will occur more frequently than can be expected by chance) can be tested by comparing observed with expected frequencies of coalitions. In the 144 games played, the weak man was included in 118 coalitions and excluded from 26. Since, however, under the S-S contention condition a weak-strong coalition was forced (although in two cases the two strong players did coalesce), only games played under the no S-S contention condition are pertinent to hypothesis 1. In the 72 games, the weak man entered 48 coalitions and was excluded from 24. These are precisely the frequencies expected by chance. Thus the evidence for hypothesis 1 is negative.⁶

Hypothesis 2 asserts that in weak-strong coalitions, the weak man will obtain less than 50 per cent of the prize. Of the 118 coalitions including the weak man, in 75 he received less than 50 per cent of the prize and in 43 he received more than 50 per cent. Since these trials are not independent, the X^2 test cannot be applied to these data. If, however, data for the twelve triads under the various experimental conditions are considered separately, as in Table 2, then exact probabilities for the occurrence of frequencies over and under 50 per cent can be computed on the assumption that splits above and below 50 per cent are equally likely. It will be noted, from Table 2, that in nine of the twelve experimental conditions, the splits accord with the hypothesis. Two exceptions occur under the S-S-contention-and-no-W-S-contention condition; here we expect the weak man to obtain larger proportions of the prize (hypothesis 4). The closest approximations to completely chance distributions, over all cases, occur in the higher weight conditions, where

⁵ It was initially intended to replicate the Latin Square four times. For reasons to be indicated, this intention was abandoned. The present analysis is based on the results of one Square. Thus, order and triad differences are confounded. Our present concerns do not, however, require that these be separated. Further, these data represent the throwing together of triads which initially were to be parts of independent Latin Squares, one a "Tuesday" group, the other a "Thursday" group. Since persons were assigned to triads randomly, and since there is no reason to suspect that "Tuesday" and "Thursday" groups of volunteer subjects differ, this treatment seems justified.

⁶ Although irrelevant to this hypothesis, the frequencies with which the weak man entered coalitions were identical under the two W-S contention conditions.

TABLE 2
*Proportions of the Prize Obtained by Weak Man (W-S Coalitions Only),
 by Contention Conditions and Weights*

Proportion of the prize	No S-S Contention						S-S Contention						Total
	No W-S Contention			W-S Contention			No W-S Contention			W-S Contention			
	Weights			Weights			Weights			Weights			
	1	3	5	1	3	5	1	3	5	1	3	5	
	1	3	5	1	3	5	1	3	5	1	3	5	
1-49	7	5	5	7	8	5	7	4	2	10	9	6	75
50-over	2	2	3	1	1	2	4	8	10	2	2	6	43
p *	.09	.23	.36	.04	.02	.23	.27	.19	.02	.02	.03	.61	118

* Probability of obtaining a split at least this unbalanced if the true probability is .5.

again we expect the weak man to obtain higher proportions of the prize (hypothesis 3a). We conclude that, though the evidence is equivocal, in general it is in accord with hypothesis 2.

Table 3, the Latin Square analysis, has been arranged to order horizontally the 12 experimental conditions and to order vertically the 12 triads. Cell entries are the proportion of the prize obtained by the weak man. Zero entries represent games in which the weak man did not enter the coalition formed.

Inspection of this table reveals that there are virtually no differences, as the weak man's weight varies, in the frequencies with which he enters coalitions. This is true whether the entire matrix is considered, or whether, more relevantly, only that section of the table dealing with games played under the no S-S contention condition is considered. Thus, the evidence for hypothesis 3a (that increases in weight of the weak man relative to the two stronger others will result in more coalitions including the weak man) is negative.

With respect to hypothesis 3b (that increases in weight of the weak man relative to the two stronger men will result in larger proportions of the prize being obtained by the weak man) the F ratio for weights in Table 3 is not significant, although it approaches the .05 level. This analysis includes zero scores, i.e., both entering a coalition and achieving a proportion of the prize are taken into account. Since we are interested in how well the weak man does, failing to enter a coalition is relevant and this treatment of the data seems justified. One can, however, ask another question: in those coalitions in which the weak man is included, how well does he do? In what follows, both questions will be treated.

When the divisions of the prize are trichotomized so that N's in the three categories are as similar as possible, it appears that the frequency with which the weak man obtains larger proportions of the prize does increase with weights. This holds whether all cases are considered, or only those in which

TABLE 3

Analysis of Variance, Latin Square (Cell Entries: Proportion of Prize Obtained by Weak Man)

Triad	No S-S Contention						S-S Contention					
	No W-S Contention			W-S Contention			No W-S Contention			W-S Contention		
	Weights			Weights			Weights			Weights		
	1	3	5	1	3	5	1	3	5	1	3	5
1	49	0	49	0	50	40	100	98	75	49	40	40
2	0	25	0	0	0	45	20	33	45	33	0	45
3	15	0	45	15	1	0	15	99	50	20	50	50
4	15	0	0	30	25	67	20	25	50	15	30	50
5	40	0	0	0	40	50	95	50	95	30	40	50
6	15	0	40	15	40	40	15	25	50	15	30	35
7	0	40	50	1	0	0	50	70	60	50	35	50
8	20	33	20	0	0	0	25	50	40	15	33	40
9	55	50	0	40	40	25	40	85	50	40	40	50
10	1	15	1	20	25	45	85	99	60	25	35	30
11	50	50	50	60	40	0	40	5	50	60	50	50
12	0	33	50	25	33	0	0	50	50	25	33	40

Completed Analysis of Variance

Source	Sum of Squares	df	Mean Square	F
Triads	10524.19	11	956.74	2.24
Sequence	4625.17	11	147.74	.35
Experimental conditions				
A. S-S contention condition	17095.70	1	17095.70	39.98
B. W-S contention condition	2062.38	1	2062.38	4.82
C. Weight conditions	2390.87	2	1195.44	2.80
Interactions				
A × B	2078.00	1	2078.00	4.86
A × C	362.66	2	181.33	.43
B × C	78.68	2	39.34	.09
A × B × C	764.57	2	382.29	.89
Residual	47041.22	110	427.65	
Total	86923.44	143		

$$F_{.05} [1,110] = 3.93$$

$$F_{.05} [2,110] = 3.08$$

$$F_{.05} [11,110] = 1.87$$

weak-strong coalitions were formed. In the former, when the weak man carried the weight 1, he received between 0 and 24 per cent of the prize in 52 per cent of the games played; carrying weight 3, he received 25 to 48 per cent of the prize in 48 per cent of the games; and carrying weight 5, he received 49 per cent or more of the prize in 46 per cent of the games. If only games in which weak-strong coalitions were formed are considered, when the

weak man carried the weight 1, he received between 1 and 32 per cent of the prize in 58 per cent of such games; carrying weight 3, he received 33 to 49 per cent of the prize in 41 per cent of the games; and carrying weight 5, he received 50 per cent or more of the prize in 5 per cent of the games.⁷

If comparisons are made between scores for the same subjects under various weight conditions, and the X^2 test for correlated proportions as in Table 4 is used, the data substantiate the impressions that changes in proportions as weights increase are in an upward direction.⁸

TABLE 4

Changes in Proportions of Prize Obtained by Weak Man by Changes in Weights, Correlated Samples

To Weight 3				To Weight 3			
		33-over	1-32			50-over	1-49
From	1-32	11	8	From	1-49	4	20
weight 1	33-over	13	1	weight 1	50-over	7	3
		$X^2=8.3$				$X^2=.14$	
		$p<.01$				$.80>p>.70$	
To Weight 5				To Weight 5			
		33-over	1-32			50-over	1-49
From	1-32	7	1	From	1-49	11	11
weight 3	33-over	23	3	weight 3	50-over	10	2
		$X^2=1.6$				$X^2=6.2$	
		$.30>p>.20$				$.02>p>.01$	
To Weight 5				To Weight 5			
		33-over	1-32			50-over	1-49
From	1-32	16	3	From	1-49	10	16
weight 1	33-over	13	1	weight 1	50-over	7	0
		$X^2=13.2$				$X^2=10.0$	
		$p<.01$				$p<.01$	

In Table 5, the results of "t" tests for the differences between means for correlated samples, performed on the pairs of columns of Table 5, are presented.⁹ The direction of the changes in mean scores as weights increase

⁷ Since the data in the three categories of these comparisons are not independent, no statistical test of these differences could be made.

⁸ Since the results of the X^2 test when all cases were used and when only W-S coalitions were used are identical, only the latter are presented here. Both cutting points were used to avoid the possible biases inherent in selecting one, and as a means of showing the upward flow of the data. The data were dichotomized in two ways: first comparing proportions in the 1-32 per cent range with those above 32 per cent; second, comparing proportions in the 1-49 range with those above 50 per cent.

⁹ Only analyses which include zero scores are presented. To eliminate these for the no S-S contention condition would leave too few cases for a meaningful test. Eliminating the two pairs involving zero scores in the S-S contention condition does not materially alter the findings.

is in accord with the hypothesis in 10 of the 12 comparisons. However, only in three cases are these differences statistically significant.

In summary, while the evidence is not clear cut, there is considerable support for hypothesis 3b, i.e., when the weak man is involved in the coalition, he tends to receive larger proportions of the prize as his weight increases.

Hypothesis 4 states that contention between the two strong members of the triad will result in larger proportions of the prize being obtained by the

TABLE 5

Mean Proportion of Prize Obtained by Weak Man for Weight Conditions and Tests of Differences between Means by Contention Conditions (All Cases)

Contention Condition	Weights of Weak Man			Comparison	df=11	p (one tailed test)
	1	3	5			
No contention	21.66	20.50	25.42	1 vs 3	.15	.90>p>.80
				3 vs 5	.58	.30>p>.25
				1 vs 5	.41	.35>p>.30
No S-S and W-S	17.17	24.50	26.00	1 vs 3	1.21	.15>p>.10
				3 vs 5	.20	.90>p>.80
				1 vs 5	.90	.20>p>.15
S-S and no W-S	42.08	57.42	56.25	1 vs 3	1.51	.10>p>.05
				3 vs 5	.13	.90>p>.80
				1 vs 5	2.13	.05>p>.02
S-S and W-S	31.42	34.67	44.17	1 vs 3	.65	.30>p>.25
				3 vs 5	2.49	.02>p>.01
				1 vs 5	3.07	.01>p

weak man. The F ratio for the S-S contention condition, in Table 3, is highly significant. In part, this reflects the fact that the bulk of the coalitions not including the weak man occurred in the no S-S contention games. When only weak-strong coalitions are examined, the distributions of the proportion of the prize received by the weak man under the two S-S contention conditions are in accord with the hypothesis: in the no S-S contention games, he receives less than 32 per cent of the prize in 40 per cent of the games, 50 per cent or more of the prize in 23 per cent of the games; under the S-S contention condition, he receives less than 32 per cent of the prize in 26 per cent of the games, 50 per cent or more of the prize in 46 per cent of the games. Impressions gathered from these distributions are reinforced by the data of Table 6, containing X^2 analyses of changes in scores for correlated samples, and by Table 7, which includes tests of the significance of the differences between the mean value of the prize for matched pairs of cases, and includes only pairs in which the weak man entered a coalition in both games. Although the direc-

TABLE 6

Changes in Proportion of Prize by Changes in Contention Conditions, Correlated Samples

	To S-S Contention				To S-S Contention		
		50-over	1-49			33-over	1-32
From no S-S	1-49	15	22	From no S-S	1-32	9	10
contention	50-over	7	4	contention	33-over	26	3
		$X^2=6.37$	$.02 > p > .01$			$X^2=3.00$	$.10 > p > .05$

tion of the difference is that expected under both the W-S and no W-S contention conditions, only in the latter is the difference statistically significant. We conclude, however, that hypothesis 4 is supported by the evidence, since we expect, on the basis of hypothesis 5, that W-S contention will reduce the rewards obtained by the weak man through restricting his bargaining power.

Table 3 contains evidence supporting hypothesis 5, in that the F ratio for the relevant W-S contention condition is significant. Since the cells containing zero entries are equally divided between the W-S contention conditions, there need be no qualification of the results on this basis. A qualification is introduced, however, by a consideration of the significant F ratio for the first order interaction of contention conditions. The significance of this interaction phenomenon can best be seen in the data of Table 8, where results of "t" tests for differences between means of the W-S contention conditions are presented, treating separately the cases by S-S contention conditions. Analysis of actual frequencies supports this conclusion. The mean scores for the W-S contention conditions, when the no S-S contention condition applies, are almost identical. Excluding pairs involving zero scores does not alter this conclusion. On the other hand, the differences for the W-S contention conditions when the S-S contention condition applies are highly significant. This general conclusion holds when data are analyzed treating each weight separately. Conclusions about hypothesis 5 must be tempered by this observation of interaction: absence of constraint on the weak man operates to his benefit only when the strong pair are in contention with one another.

TABLE 7

Mean Proportion of Prize for S-S Contention Conditions and Tests of Differences between Means, by W-S Contention Conditions (W-S Coalitions Only)

W-S Contention Condition	Mean Scores, S-S Contention Conditions		t	df	p (one tailed test)
	No S-S Contention	S-S Contention			
No W-S contention	33.79	52.58	3.03	23	$p < .01$
W-S contention	34.65	38.39	1.06	22	$.20 > p > .10$

TABLE 8

Mean Proportion of Prize for W-S Contention Conditions and Tests of Differences between Means, by S-S Contention Conditions (All Cases)

Condition	Mean Scores, S-S Contention Conditions		t	df	p (one-tailed test)
	No W-S Contention	W-S Contention			
No S-S contention	22.53	22.56	.01	35	p>.90
S-S contention	51.92	36.75	3.40	35	p<.01

DISCUSSION

The theoretical issues which motivated this research were diverse. Those for which the evidence is most clear pertain to Simmel's notions concerning *tertius's* advantage when contention exists between the two stronger members and the reduction in *tertius's* advantage when his freedom to choose between potential partners is restricted. In general, Simmel is supported by the data of this study.

With regard to the effect of increases in weights of the weak man on the share of the prize he obtains, different theoretical bases provided alternative hypotheses. Following Simmel, one expects that increases in relative weight are of no significance. Following the implications of Capiow's formulation and Vinacke and Arkoff's research, one expects that such increases will result in increasing rewards. The data support the latter view, deriving from what we have called perception theory. This, however, is true only for divisions of the prize and not for gaining access to coalitions.

A third issue involves the contrast between predictions made on the basis of game theory and on the basis of the Vinacke-Arkoff interpretation of their results. In contrast to the Vinacke-Arkoff finding that the most frequent coalitions were either AB or AC, our results are that the weak man entered coalitions no more frequently than would be expected by chance.¹⁰ This chance distribution of persons involved in coalitions is that which would be predicted by game theory, i.e., if participants bargained in terms of their best strategy to maximize winnings. We may ask, therefore, whether the frequencies found in this study are evidence for the game theory view. Or, to reverse the question, if coalitions did disproportionately include the weak

¹⁰ In our procedure, coalitions were forced. However, Vinacke and Arkoff's subjects were allowed to play without forming coalitions if they so chose. If the eleven games in triad Type 3 which Vinacke and Arkoff report involved no coalition are interpreted as the equivalent of a BC, or, in our terms, a strong-strong coalition, since the two equally strong players would receive equal shares of the prize, then their results would also show that A, the weak man, was no more frequently involved in a coalition than would be expected by chance.

man, would this be evidence that the participants were not behaving "rationally?"

Before we attempt to answer these questions, another may be posed. As Vinacke and Arkoff interpret Caplow, they assert that his theory predicts that divisions of the prize will reflect the initial strength of the players. From game theory, the expectation is that divisions will be 50-50. Evidence from both the Vinacke and Arkoff and the present study indicates that the prize obtained by the weak man, more frequently than could be obtained by chance, is less than 50 per cent. Is this evidence that players do not play these games "rationally"?

When the present research was begun, we were confident that the data obtained through the procedures described could answer these questions. At this point we are certain that, without modification, they cannot.¹¹ The grounds for this disillusionment may have occurred to the reader. The rationales for predictions made by Vinacke and Arkoff are based on assumptions concerning the process by which players make decisions and bargain for coalitions. Outcome predictions are made on the basis of these process rationales and demonstration of a given outcome is used as evidence that a given process does operate. In short, the inference is that if outcomes (in the form of frequencies of coalitions and division of prizes) correspond to what would be predicted if a given process had taken place, then that process has taken place.

Yet it is obvious that such inference is extremely risky unless one outcome alone is logically and empirically possible when and only when a given process operates. The fact that a weak man does not enter more coalitions than would be expected by chance could presumably be attributed to the rationality of the players. But chance and rationality are not the same, and there is no way of determining from outcome data alone which process was, in fact, operating. Data concerning process as well as outcome are needed.

Such data can be provided by observing the interaction of participants as they bargain for divisions of the prize and coalitions. It would then be possible to determine whether "rational" bargaining produced a 50-50 prize division and, over a series of triads, whether the frequency of inclusion of the weak man in coalitions, if no different from chance expectations, was the result of "rational" bargaining behavior. The bargaining for coalitions in terms of promises to share the prize would provide crucial data for the experimental game used in this study. We paid careful attention, therefore, to the verbal exchanges among players but found the experimental situation was not as conducive to competitive bargaining as we expected. Often, initial

¹¹ The inadequacy of the procedures for settling this issue in no way reflects on their adequacy for answering other questions raised in this study.

offers by one player to another were accepted without the third person's entering into the bargaining at all.

In some games played by some triads, players tossed the die to determine who would coalesce with whom; some even went so far as to toss the die to determine the division of the prize among the partners. In other cases, a player who saw the various possibilities for coalition formation coached his "opponents" as to their best strategy even though this could, and sometimes did, result in his being left out. The actual altruism, consideration and empathy that existed in this situation would have been overlooked and the outcome instead attributed to competitive, rational processes if outcome data alone were examined. In other triads, persons evolved norms with regard to which opponents would join under which power weight and contention condition and what a "fair" split of the prize would be. These norms were adhered to whenever other games of similar contention conditions were played. Cases such as the ones mentioned in this paragraph were not included in the data analyzed for the present paper. (They did, however, indicate that an adequate test of game theory versus perception theory is not possible using this experimental game and setting.)

Moreover, a game theory outcome is most likely only under a rare combination of conditions such that all three players are completely and equally "rational." Interestingly enough, on some such occasions which did emerge, the players, in recognition of the fact that bargaining would be interminable since each would eventually agree to join any other for an equal division of the prize, would toss a die to determine coalitions. Here is a rational selection of a chance mechanism for determining coalition partners. If only one or two players operate with a game theory perspective, an offer to give greater than half the prize made by the third person may be quickly accepted. This would be a "rational" decision although the outcome was not a 50-50 split. It is possible that it was in cases such as these that the "rational" player(s), because of the absence of a strong competitive motivation, felt a moral obligation to reveal his insights to the other players as mentioned previously. We have also found that it is possible for a single player, if he bargains quickly and forcefully and counters B's offer to C by a better offer to B for example, to produce 50-50 prize division outcomes although he and he alone accurately perceives game theory strategy. We did not observe any such cases.

This discussion, dealing with the possibility of testing game theory predictions, suggests that no reliable inferences concerning process can be made from outcome data alone. Furthermore, the experimental situation and game used by Vinacke and Arkoff are not adequate for testing hypotheses based on game theory assumptions as to rational, competitive bargaining processes.

In current work on problems discussed in this paper, some modifications in

research procedure have been introduced. In the attempt to illuminate the process by which coalition decisions are reached, a pre-game questionnaire is used, and written messages of offers, rejections and acceptances required. The pre-game questionnaire asks a subject to indicate what he would offer the other players, and what they would offer him, if such offers actually were to materialize. It also asks a subject to indicate what offers he would make if he were in the positions of the other participants. We expect that the pre-game data, and the record of sequence of offers, etc., made possible through the written messages, will throw light on the meaning of the outcome data. Further, we are now using a screen to separate players physically in an attempt to minimize the intrusion of variables extraneous to the immediate concerns of the research. True, from the standpoint of a theoretical social psychology, this eliminates variables of some importance; for example, face-to-face interaction. In terms of research strategy, these variables must either be systematically studied or controlled; it cannot be assumed, as in the earlier research, that they can be neglected. In the interest of developing a "base-line" for coalitions and prize divisions, we have chosen to exclude them insofar as possible. It is our intention, in future studies, to introduce such variables systematically.

Manuscript received: May 18, 1959

Revised Manuscript received: November 4, 1959

Sheldon Stryker

Department of Sociology

Indiana University

Bloomington, Indiana

REFERENCES

1. Caplow, T., "Further Developments of a Theory of Coalitions in the Triad," *American Journal of Sociology*, 1959, 64, 488-493.
2. Caplow, T., "A Theory of Coalitions in the Triad," *American Sociological Review*, 1956, 21, 489-493.
3. Simmel, G., *The Sociology of George Simmel*, translated and edited by K. Wolff, Glencoe, Ill.: The Free Press, 1950.
4. Vinacke, W. E., "Sex Roles in a Three-Person Game," *Sociometry*, 1959, 22, 343-360.
5. Vinacke, W. E., and A. Arkoff, "An Experimental Study of Coalitions in the Triad," *American Sociological Review*, 1957, 22, 406-414.

Coalitions in the Triad: Critique and Experiment

HAROLD H. KELLEY AND A. JOHN ARROWOOD,

*Laboratory for Research in Social Relations, University of Minnesota*¹

Vinacke and Arkoff (4) have presented an experiment which tests some of Caplow's hypotheses (2) about how the relative power of three persons affects the formation of pair coalitions. The situation studied is one in which each person is trying to obtain for himself as much of some valuable but scarce commodity as he can. The three individuals differ in ways relevant to their ability to gain a share of the rewards (referred to as "power"). This work has been especially interesting because of the paradoxical result that with certain distributions of power among the three individuals, in Caplow's words, "the triadic situation often favors the weaker over the strong." Under certain conditions the stronger of the three is at a disadvantage and actually receives the smallest share of the available rewards.

The purpose of the present research is to state with greater precision than heretofore the conditions under which this phenomenon prevails and to test experimentally some of the limits of these conditions. In order to do so, it is necessary to clarify certain ambiguities in the concept of power as used in the Vinacke and Arkoff experiment. This clarification is accomplished through use of the analysis of power provided by Thibaut and Kelley (3).

The problem can best be illustrated by a brief description of the Vinacke and Arkoff procedure. Three subjects play a game in which each moves his counter along the spaces of a game board. The first one to reach the goal receives a prize of 100 points. On successive trials, the experimenter rolls a single die and each player advances a number of spaces determined by the product of two numbers: (a) the number of pips turned up on the die and (b) a "weight," ranging from 1 to 4, which was randomly assigned him at the beginning of the game. For example, in one game player A may have weight 4, player B, weight 2, and player C, weight 1. Since all players start at the same point on the board and move each time the die is cast, the person assigned the largest weight automatically wins. A further rule, however, enables any pair of players to form a coalition by combining their weights at any time during the game. When they do so, they are given a single counter placed at a position equal to the sum of the distances the two have attained at that time. On subsequent rolls, they advance according to the sum of

¹ This research was supported by grant G-5553 to the senior author from the National Science Foundation.

their two weights. The formation of a coalition is acknowledged by the experimenter only when the two players have agreed upon how they will divide the 100 point prize, should they receive it; and, once formed, a coalition is indissoluble for the remainder of that game. Thus, the individual or coalition that can mobilize the largest weight automatically wins that game and there is really no need for going through the motions of rolling the die.

The weight each player receives is said to constitute his *power*, but consider this point more closely. In what sense does a player with a weight of 4 have more power than a player with a weight of 2? In the game where the three weights are 4, 2 and 1, the player with the 4 weight has power in the sense that he is able, regardless of the actions of the other two players, to induce his "environment" (the game) to give him the prize. However, in the game with weights 4, 3 and 2, the player with 4 weight can exercise this control over the environment only if the other two players fail to form a coalition. Since any pair can mobilize more weight than the remaining person, each pair has the same amount of power over the third person as does any other pair. The variability in 4's outcomes is as much under the control of the joint actions of the other two players as is the variability in either 2 or 3's outcomes. Hence, as Vinacke and Arkoff point out, the initial weights in the 4-3-2 game are irrelevant with respect to the power a person has in the three-way bargaining situation.

In view of this logical analysis of the objective interdependency relations among the three players, Vinacke and Arkoff's results from the 4-3-2 and similar games are unexpected: The three players treat the weight 4 as if it does yield greater power. As his price for entering a coalition, player 4 apparently asks for a lion's share of the prize, because he typically receives more than 50 of the 100 points when he is in a coalition. Furthermore, players 2 and 3 tend to form the majority of the coalitions, presumably because each one can make a better deal with the other than he can with 4. From the point of view of rational analysis, then, the subjects act inappropriately, attributing to 4 a power that he does not in fact possess. The irony of the situation is that this erroneous belief about 4's advantage, which he usually shares, works to his disadvantage in the long run because of his exclusion from coalitions.

The first of the experiments reported below has the purpose of showing that Vinacke and Arkoff's data are, in a sense, spurious. They reflect a misunderstanding of the experimental situation that is not intrinsic to it, but results from the complexity of their total procedure. Confronted with the complexity, subjects erroneously equate initial weights with real power. (Our reasons for believing this erroneous assumption to be a reasonable one, considering the circumstances, are given in the Discussion.) The experimental

hypothesis is that with a simpler procedure, subjects will acquire an adequate understanding of the true power relations and act more in accord with a rational analysis of the situation than the Vinacke and Arkoff data would suggest.

The second experiment has the purpose of testing Caplow's hypothesis under conditions where the power differences among the three persons are real rather than illusory. We expect the resulting bias in coalition formation (predominantly between the weaker members) to persist even though subjects are permitted thoroughly to familiarize themselves with the situation. (The limiting conditions under which the Caplow effect can be expected to appear are also described.)

EXPERIMENT I

In the Vinacke and Arkoff procedure, the relationship between weights and power is quite complicated. Each trio of subjects is required to play a series of games in which six different sets of weights are used. In the games played with a given set of values, a player need not have the same weight twice. With some sets of weights (such as 4-2-1) initial weight is relevant to power, and with others (such as 4-3-2) it is not.

In the present experiment, in order to simplify the subjects' task, only one set of weights was used (4-3-2) and each triad was given a lengthy series of trials, each player keeping the same weight throughout. It was expected that, with repeated experience in the single situation, the subjects' analyses of it would come to correspond with the analysis presented above, and their coalition formation and bargaining behavior would be increasingly consistent with an understanding that all players in the game have the same power.

Procedure

Ninety male students, volunteers from an introductory (sophomore) psychology class, served as subjects in thirty experimental triads. There is no reason to believe that the subjects were notably different from Vinacke and Arkoff's. Data were gathered first from 20 triads and, later, from ten additional triads. The first series were given a variable number of trials (10 to 70 trials, 26 on the average), and in the second series all groups completed 20 trials.

With the exceptions noted above, the experimental procedure followed that of Vinacke and Arkoff as closely as was possible from the available statements of their procedure. A major difference is that, whereas they gave very brief introductory instructions and relied on informal answers to ques-

tions to clarify the procedure, we gave rather full formal instructions and tried to minimize the necessity for subjects to ask questions. The subjects were given an individualistic orientation, it being emphasized that each was to accumulate as many points for himself as possible, attempting to maximize his outcomes without regard to those of any other players.

Results

In Table 1 are presented the frequencies of occurrence of the various coalitions. Data from our first and last three trials are presented for comparison

TABLE 1
Frequency of Occurrence of Various Coalitions in the Three Experiments

Coalitions	Vinacke and Arkoff's Three Trials		Experiment I						Experiment II			
			First 3 Trials		Last 3 Trials		(Coalitions)	First 3 Trials		Last 3 Trials		
			N	%	N	%		N	%	N	%	
2-3	59	66	41	46	37	41	(0-2)	24	53	29	64	
2-4	20	22	24	27	26	29	(0-4)	12	27	13	29	
3-4	9	10	21	23	27	30	(2-4)	9	20	3	7	
No coalition	2	2	4	4	0	..		0	..	0	..	
Totals	90	100	90	100	90	100		45	100	45	100	
<div> $\chi^2_{n=88}=47.07$ $2df$ $p<.001$ </div> <div> $\chi^2_{n=90}=2.47$ $2df$ $.20<p<.30$ </div> <div> $\chi^2_{n=45}=22.93$ $2df$ $p<.001$ </div>												
<div> $\chi^2_{n=86}=8.12$ $2df$ $.01<p<.02$ </div> <div> $\chi^2_{n=45}=8.40$ $2df$ $.01<p<.02$ </div>												

with those from Vinacke and Arkoff's three trials in the 4-3-2 condition. They calculated a Chi-square for the 88 instances of coalition formation (excluding the two instances of "no coalition") to determine the likelihood of the departure of the observed distribution from a theoretical distribution in which the three possible coalitions occur equally often. This procedure is not strictly justified, inasmuch as each group of subjects provides three instances of coalition formation, hence the various entries are not independent. However, we present similar Chi-squares in order to provide some basis for comparing the two sets of results.

It appears that the present procedure yields a distribution of coalitions which, although biased in a manner similar to Vinacke and Arkoff's, is closer to the chance distribution than is theirs. The divergence from the earlier experiment appears even on the first three trials. A comparison of our first three with theirs yields a Chi-square of 8.48, $p = .02$. (The distribution of first coalitions formed in each triad resembles their distribution most

closely, the percentages, in order, being 63, 20 and 17.) These results suggest that, either because of the concentrated experience with the single situation, or perhaps because of greater clarity in the experimental instructions, a number of subjects became aware of the objective power relationships rather early in the game.

From the first to the last trials there is not a significant change in the incidence of the three coalitions, hence we cannot be sure that further experience increased their understanding of the situation. However, this result must be interpreted in the light of the rather limited degree to which the occurrence of the early coalitions departed from chance. To the extent that subjects achieved an understanding of their relationships very early, further improvement in this respect was limited. Upon examination of the data on a trial-by-trial basis, it is clear that the learning is very rapid, so that after the first three or four trials there is little more than chance exclusion of 4 from coalitions.

In Vinacke and Arkoff's experiment, associated with the tendency for 2 and 3 to be reluctant about forming coalitions with 4, was the tendency for him to receive more than half of the 100 point prize when he did manage to enter a coalition. This effect appears in the early trials of the present experiment and declines, though not significantly so, during its course. During the first three trials, player 4 was a member of a coalition in 26 triads and in 17 of these he managed to come out with more than 50 per cent of the rewards. During the last three trials of the experiment, this was true in only 10 of 29 possible instances. Of the 25 groups in which person 4 was in coalitions during both the first and last three trials, his share of the coalition reward declined in 13 instances, increased in six instances and did not change in the remaining six. Omitting the last six, the difference between the number of decreases and increases yields a p -value of .168 by the Sign test.

There are, then, two behavioral manifestations of the players' perception that 4 has the most power: (a) He is excluded from coalitions and, (b) when he is included in coalitions, he receives more than half the points. When early and later trials are compared, the extent to which 4's frequency of being included increases and the extent to which his percentage winnings per trial decrease provide two indicators of downgrading in the perception of his power. A comparison of the first and last three trials of the experiment reveals that in 18 triads the 4 man was downgraded on both of these indicators or downgraded on one with no change on the other; in six triads, he was upgraded on both indicators or upgraded on one with no change on the other; and in six triads, he did not change on either indicator or was upgraded on one and downgraded on the other. When the last six are omitted from the analysis, the Sign test indicates there was significantly more downgrading than upgrading ($p = .026$).

On the questionnaire at the end of the experiment all subjects were asked the following open-ended question: "Many subjects believe that 4 is more to be feared, has greater potential, etc. Did you at any time think that this was the case?" All but 14 of the 90 subjects admitted that they at some time had held this belief. This estimate of the extent of the belief, if it errs at all, is probably an underestimate because of the general realization at the end of the experiment of the incorrectness of this view. Most of those admitting to this belief reported having it "at first" or on the first one or two trials. Subjects with the three different weights were equally susceptible to this belief.

The next question asked when the subject realized that no position has any more power than any other and that nobody is justified in asking for more than half the 100 points. Forty-three of the 90 claimed to have realized this before or during the first three trials. Another 33 did not localize their insight so sharply in time. Only 13 admitted they had never realized this. On the basis of their answers to other questions about power relations, preferred position, position likely to win, and ease of bargaining with various positions, another nine subjects were added to the latter category of those who had failed to attain a correct understanding of the situation by the end of the experiment.

In brief, the self-report data suggest that, while some 85 per cent of our subjects believed at some time during the experiment (largely, in the early stages) that the 4 weight carries greatest power, only 25 per cent held this belief at the end of the experiment. As far as we could judge from their answers to questions asked after the experiment, the other 75 per cent had achieved a correct understanding of the power relations and apparently most of them did so during the actual trials of the experiment.

One may ask to what extent this change in belief about power relations reflects the subjects' direct experience with the game, as opposed to their being taught by the small number of their colleagues who had analyzed the situation correctly from the start. To answer this question, a comparison was made between those groups in which, from the sound recordings of the discussion, there appeared to have been some possibility of "teaching," and the remaining groups. There were no differences between the two sets of data either in the amount of learning that took place or in the coalitions formed. Hence, explicit teaching seems not to account for the observed effects.

A final open-ended question asked: "Why do you think that many subjects would believe that 4 is more to be feared, has greater potential, etc.? What is there in the situation that leads to this belief?" The most frequent response, given by 34 of the 90 subjects, dealt with the fact that 4 would win invariably

if no coalitions were formed. Another 13 merely stated, without further amplification, that 4 was the largest number. Nine subjects mentioned the multiplicative aspect of the game, pointing out that multiples of 4 are larger than multiples of 2 or 3. Apparently, the 56 responses in these three categories either discount or overlook the possibility of coalition formation. A fourth category of response deals with 4's "psychological" or "cultural" power. Seventeen of the 90 suggested that it was natural to react to 4's higher weight in the light of their previous experience in games that higher numbers are generally better ones and in everyday life that quantity often has the upper hand. This answer is silent concerning the possibility of coalitions but highlights the stereotype that "more" is "better."

It is difficult to evaluate the validity of these post-experiment explanations for the earlier misinterpretation. Because the misperception of 4's power is largely corrected by the end of the experiment, subjects may be somewhat reluctant or even unable to discuss the real basis for their mistaken views. Hence the reasons given by the 22 players who seem never to have realized the true nature of the power situation may be especially valuable indications of the source of the error. Seven of the 22 seemed to believe that coalitions including 4 were somehow more sure of obtaining the prize than were other coalitions or that 4 was the only player capable of bargaining and that the other players had either to accept his terms or receive nothing. We believe it noteworthy that it is only among these 22 players that we find the assertion (made by another seven of them) that a high-weight player is justified in demanding a majority share of the coalition reward because he *contributes more* to the coalition. One might guess that this interpretation figures more prominently in the early reaction to the situation than the overall figures would indicate. The comments about 4's larger number and the multiplicative aspects of the game may well be oblique references to earlier beliefs (which now appear to the subjects as totally unjustified) that the higher weight player makes a greater contribution to coalition success.

EXPERIMENT II

Under the Vinacke and Arkoff procedure, the alternative to being in a coalition on any given trial has the same value (zero) for each player regardless of his weight. Hence, 4's outcomes are as much subject to control through the joint actions of the other two players as are either 2 or 3's outcomes. It is in this sense that 4 has no more power than they. In the present experiment, real power differences are created by giving the three subjects differential ability to obtain rewards from the game, an ability that can not be attenuated by the actions of the other two persons. This is done by giving each person a specific alternative level of outcomes which he receives if he

fails to gain membership in a coalition or if, once in a coalition, he and his partner fail to reach agreement on a division of the spoils. (The bargaining involved in the division of the prize *follows* rather than precedes the choice of coalition partners.) The person with a higher alternative value has high power in the sense that (a) he is less dependent upon getting into a coalition, and (b), during the bargaining following coalition formation, he can hold out for a larger share because he has less to lose if no agreement is reached. The long run effects of the latter, demanding a preponderant share of the reward, are of rather little concern to him because of the first fact. At the same time, factor (b) makes him less desirable than a weaker person as a coalition partner. Hence, we would expect that the poorer a player's alternative, the greater the likelihood of his being included in a pair coalition. Of course, this will be true only when the size of prize given a coalition does *not* increase in proportion to the power of its members. (In the present case, as in the Vinacke and Arkoff procedure, the prize is the same for all coalitions.) It is under these circumstances that Caplow's statement is relevant—that "the weakest member of the triad has a definite advantage, being sure to be included in whatever coalition is formed."

Procedure

Forty-five male students, volunteers from an introductory (sophomore) psychology class, served as subjects in 15 experimental triads. The task was presented as a simple business game in which each player was a corporation chairman, controlling a certain share of the market each month. Each player's object was to accumulate as many points for himself as possible, not to compete but to attempt to maximize his own outcomes without regard to how this might affect the outcomes of the other players. Each subject was randomly assigned a weight—either 4, 2, or 0—which represented the number of points he could earn on each trial if he chose to play the game independently. Any pair of subjects, however, had the option of forming a coalition which was then given one minute to decide in what manner to divide a ten-point prize between them. Coalitions were formed by a series of written choices. At the beginning of each trial, each subject privately indicated the number of the other player with whom he would most like to form a coalition. Reciprocated choices became coalition partners, thereby having the opportunity to attempt to decide how they wanted to divide the ten points between them. The third man, the player *not* in the coalition, was paid off immediately with the number of points equal to his weight or alternative and did not enter into the bargaining for that trial. (If there happened to be no reciprocated choices on a given trial, the subjects were requested to consider the problem again and indicate their choices once more. This pro-

cedure was continued until a reciprocated choice appeared. This was necessary on only 43 of the 300 trials. Most of the instances of non-reciprocation were found early in the game.) If the two members of the coalition reached some mutually satisfactory division of the ten points during the minute allotted for bargaining, they then received that number of points as their scores for that trial. If, however, they did not agree before the time limit, they forfeited the ten points and each received the number of points equal to their weights or alternatives. A time limit was placed on the bargaining so that the weaker player could not gain power by stalling and controlling time. Each subject retained the same weight throughout the game, and each triad completed 20 trials.

Although the weights employed in the present study differ from those used in the Vinacke-Arkoff experiment, they are comparable in at least two ways: first, if no coalitions are formed, or if no agreement is reached in the coalitions which are formed, the high-weight man will always win; and, second, the coalition is always assured a chance at a larger number of points than is any independent player. For certain analyses, then, we shall consider the present 0-2, 0-4, and 2-4 coalitions as equivalent, respectively, to the 2-3, 2-4, and 3-4 coalitions in the previous experiment.

Results

Table 1 presents the frequencies of occurrence of the various coalitions. Data from the first and last three trials of Experiment II may be compared with those from Vinacke and Arkoff and Experiment I. The relative incidence of the various coalitions for the first three trials of Experiment II does not differ significantly from that of Experiment I. The difference between the two distributions for the last three trials of each experiment is significant at the .01 level with a Chi-square of 10.72. The difference between the two distributions for Experiment II is not significant (Chi-square = 3.48, $p < .20$).

It appears, then, that the present procedure yields a distribution of coalitions which, although departing initially from the Vinacke-Arkoff distribution, does not do so as markedly as does the distribution from Experiment I and comes to approximate the Vinacke-Arkoff distribution far more closely as the trials progress.

The data also suggest that this is not merely an illusory effect which disappears with repeated experience in the situation. Table 2 presents the mean frequencies per triad of the various coalitions, and the significance levels for these differences. The 0-2 coalition occurs significantly more frequently than do either of the other possible pairs, both over the whole series of 20 trials and in both the first and second halves of the experiment; but only over

the entire series of trials does the difference between the relative incidence of the 0-4 and 0-2 coalitions approach significance. Moreover, the average frequency of 0-2 coalitions tends to increase from the first to the second half of the experiment, while that of the 2-4 coalitions tends to decrease during the same period (in both instances $.05 < p < .10$, using the ordinary *t*-test for differences).

TABLE 2

*Mean Frequency of Occurrence of Various Coalitions Overall and by Halves of Experiment II with Comparisons**

Mean Frequency per Triad	Overall	First Half of Trials	Second Half of Trials
0-2	12.33	5.67	6.67
0-4	5.20	2.80	2.40
2-4	2.47	1.53	.93
	20 trials	10 trials	10 trials
Mean Difference per Triad			
0-2 > 0-4	7.13 $t = 10.62$ $p < .02$	2.87 $t = 8.28$ $p < .02$	4.27 $t = 9.19$ $p < .02$
0-2 > 2-4	9.87 $t = 14.68$ $p < .02$	4.13 $t = 11.94$ $p < .02$	5.73 $t = 12.35$ $p < .02$
0-4 > 2-4	2.73 $t = 4.07$ $.02 < p < .10$	1.27 $t = 3.66$ $p = NS$	1.47 $t = 3.16$ $p = NS$

* The *t* values in this table were calculated using the Tukey method for multiple comparisons.

Another way of looking at the evidence is in terms of partnership choice data. Each of the positions predominantly chose the lower alternative man as a partner—i.e., 0 chose 2 66 per cent of the time, 2 chose 0 77 per cent of the time, and 4 chose 0 60 per cent of the time. Sign tests show that the results differ significantly from those expected by chance (equally frequent choice of the other two players) at beyond the .05 level in all three cases.

The previous experiments revealed an initial tendency for the 4 man to ask for and receive the majority of points from those coalitions which he did manage to enter. Over the entire series of 20 trials in the present experiment, there was a similar tendency for both 4 and 2 to receive more points than 0 in the 4-0 and 2-0 coalitions respectively (significant at better than the .02

level by the Tukey method for multiple comparisons). The slight tendency for 4 to receive more points than 2 from 2-4 coalitions proved non-significant. The over-all results are duplicated in the first half of the trials for each triad, but largely disappear during the second half—the only significant result remaining in the 4-0 coalition. As expected, on the average 4 received significantly more points per coalition from 0 than he did from 2, as did 2 from 0 over 4. The 0 man was slightly better rewarded by 2 than by 4, and especially during the second half of the game ($.05 < p < .10$). A comparison of first and second half scores reveals that 0's scores per coalition increase significantly ($p < .01$) while 4's scores per coalition decrease slightly ($.05 < p < .10$).

In a questionnaire administered at the end of the experiment the subjects were asked to indicate the following: which weight they would choose as a permanent partner; in general, which weight they believed each of the players should form coalitions with; which weight had the most power; which are the easiest and hardest weights with which to bargain; which weight would win in the long run; and which member, if either, of any coalition is justified in asking for a majority of the points. Their answers reflect the coalition formation and bargaining behavior discussed above. Only five of the 45 subjects, for example, failed to pick the available lower alternative player as a permanent coalition partner. The question: "Which of the three players has the most power?" was asked twice—once early in the questionnaire and once again at the end. At the first asking, four subjects responded that 0 had the most power, three picked 2, 36 chose 4, and two said that there were no power differences. At the later asking, 13 of the subjects who had previously singled out 4 as having power changed their answers to 2. The open-ended explanations accompanying these changes indicate a growing awareness of 2's ability to entice 0 into a coalition although still noting that, *if selected*, 4's higher weight becomes important. In general, however, subjects continue to view power as residing in a higher alternative.

DISCUSSION AND SUMMARY

It appears that Vinacke and Arkoff's procedure does initially give player 4 an illusory kind of power. In Experiment I, most of the subjects are initially subject to this misperception but apparently achieve a more correct understanding in a few trials. These results are in accord with our general contention that the phenomenon reported by Vinacke and Arkoff is limited to instances where the complexity of the learning task is so great in relation to the amount of contact and experience subjects have with it that they are not able properly to analyze it. In consequence, we witness actions that are "irrational" with respect to the analysis the experimenter makes at his

leisure. However, these actions are not necessarily irrational when viewed in the light of the understanding subjects are able to achieve under the pressures of time and task complexity. Incomplete understanding is not to be confused with irrationality.

There are at least two possible interpretations of the initial erroneous attribution of power to player 4. The first is that the initial attribution of power to 4 reflects a general pessimism about the dependability of cooperative action. Logically, player 4 is more powerful *unless* the others join forces against him. Until one knows that joint action will be instituted dependably, attributing superior power to him is not wholly unwarranted. This interpretation is suggested by the most commonly given explanation for the attribution: 4 has more power because he would win if no coalitions were formed. The declining tendency to attribute power to 4 may reflect a growing confidence that in this situation, at least, cooperative action against him is to be taken for granted.

Another possible interpretation is that our subjects have learned to use a person's potentialities in a field of independent actors as an indication of his ability to contribute to cooperative efforts. This is explicitly suggested by the comments that he makes a larger contribution to any coalition he enters and is consistent with other more general explanations provided by the subjects. In view of their likely experiences with these matters, this conclusion is a highly reasonable one. It is probably true that in everyday situations a person's effectiveness, when everyone is acting for himself, is rather closely related to how much he can add to any joint effort. Thus, the common misperception in the Vinacke and Arkoff situation may reflect a positive correlation in the social environment of the typical subject. The reader may note the similarity of this interpretation to Brunswik's explanation (1) of, for example, the size-weight illusion as reflecting a correlation between size and weight over the universe of objects the person has experienced in his physical environment. In Brunswik's terms, we are suggesting that a person's effectiveness as an individual has "ecological validity" as a cue from which to predict his ability to contribute to joint efforts, and thus enjoys considerable "impression value" or "response-eliciting power." This is the case in the Vinacke and Arkoff situation. In Experiment I, subjects initially utilize this cue extensively, but later learn that it is inappropriate in this situation and, hence, its subsequent degree of utilization declines.

The procedure of Experiment II, in contrast to that of Vinacke and Arkoff, appears to create large and lasting power differences among the members of the triad. Given differentially good alternatives to being in a coalition or, more important, to acquiescing to a coalition partner's demands, the predicted pattern of coalition formation emerges, and the weakest member of

the triad is in the most favored position when it comes to joining pair coalitions. The score and self-report data, however, suggest certain minor trends worth brief consideration. We have noted a tendency for 0's average score per coalition to increase and 4's to decrease. This finding may be taken as an indication that during the course of the trials some 0's are beginning to capitalize on their status as preferred coalition partners by asking for a larger share of the prize and some 4's are recognizing that they must be more generous in dividing the prize if they are to be allowed to enter further coalitions. We have also noted some changes in subjects' perceptions of the most powerful player and in the reasons accompanying their answers. Mention of 2's greater ability to entice 0 into coalitions suggests that some subjects are becoming aware of the truth of Caplow's hypothesis. These findings raise the interesting side problem of how a high-power individual in a situation with limited communication possibilities would go about establishing trust. Once excluded from coalitions, the high-weight man would probably tend to remain excluded, since unless he could enter a coalition there would be no way for him to demonstrate to the others that he would not use his power against them.

It must be noted that in general 4 tends to accumulate the most points during the game—93 on the average as compared with 86 for 2 and 66 for 0. However, this is probably an artifact of the relative sizes of the alternatives and the prize to be divided. By making the coalition reward larger in relation to the largest weight, one could create a situation in which the highest alternative player would, by reason of his exclusion from coalitions, end up with the smallest accumulated score. However, as the coalition prize becomes larger, the differential power implications of any given set of weights becomes less important.

One might also manipulate the weights and coalition prize in such a manner that the highest alternative player would emerge as the most preferred partner. This is an important point because it indicates the boundary conditions for the phenomena observed in Experiment II. For example, different sizes of rewards might be given to different coalitions. If the various rewards were proportional to the weights of the persons comprising the various coalitions, one would expect no difference in the relative incidence of the three possible types of coalition or even a bias in favor of coalitions including the high-power person. The latter effect would be expected, for example, if the 2-4 coalition received a prize of 12 points, the 0-4 coalition, 8 points, and the 0-2 coalition, 4 points. This would reproduce the situation where the more effective a person is as an independent actor, the more effective is the joint effort to which he contributes. It is not unreasonable to believe that, in many natural situations, joint effectiveness is a direct function (and perhaps even

a multiplicative one) of individual effectiveness. In these cases, if the above analysis is correct, coalitions would appear largely among persons of high power. On the other hand, the Caplow effect will appear when coalition effectiveness bears no relation (or a negative one) to the effectiveness of the component members.

Manuscript received: October 9, 1959

Revised Manuscript received: January 28, 1960

Harold H. Kelley

Laboratory for Research in Social Relations

University of Minnesota

Minneapolis 14, Minnesota

REFERENCES

1. Brunswik, E., *Systematic and Representative Designs of Psychological Experiments*, Berkeley: University of California Press, 1949.
2. Caplow, T., "A Theory of Coalitions in the Triad," *American Sociological Review*, 1956, 21, 489-493.
3. Thibaut, J. W., and H. H. Kelley, *The Social Psychology of Groups*, New York: Wiley and Sons, Inc., 1959, Chapters 7 and 11.
4. Vinacke, W. E., and A. Arkoff, "An Experimental Study of Coalitions in the Triad," *American Sociological Review*, 1957, 22, 406-414.

Patterns of Deviation in Work Groups

PETER M. BLAU, *University of Chicago*

It has been suggested that informal status in a group is directly associated with conformity to group norms (9). Indeed, there is considerable evidence in support of this proposition. Researchers have reported that the achievement of a position of respect and leadership depends on conformity to the values and traditions of the group (11, 13). Other studies have found that integrated members of groups are less apt than isolates to manifest deviate tendencies (7, 8, and 14, pp. 788-94). There is, however, also some contrary evidence. The experiments of Kelley and his colleagues show that integrated persons are more likely than those who do not feel fully accepted to deviate from group norms if conformity is contrary to the interest of the group (10, 5). And a field study in a government agency suggests that the more competent officials adhere less rigidly to procedures than others (2).

Several factors may account for these conflicting findings. Conformity is not all of one cloth, and neither is informal status. The significance of becoming assimilated to the basic traditions of the group might differ widely from that of rigid adherence to all minor norms. Superior rank in a group, defined by the respect and deference of others, must be distinguished from an integrated position, that is, full acceptance by others (1, 3), and these two aspects of informal status are not necessarily associated with conformity in the same way. Moreover, the influence of conformity on informal status might not be the same as that of informal status on conformity. Perhaps conformity raises an individual's status in the eyes of the rest of the group, but high status, in turn, facilitates some deviation from group norms. Indeed, Merei's findings (11) as well as those of Dittes and Kelley (5) suggest that this is the case.

The starting point of this paper is a replication, in modified form, of Kelley's experiments with laboratory groups in a field study of work groups in a public assistance agency, where group life has a saliency that defies duplication in the laboratory. It is hypothesized that integrated group members are more prone than those not fully accepted by peers to resist group pressures and hold opinions that deviate from that of the group majority. The underlying assumption is that most unintegrated persons feel constrained to improve their position and court the acceptance of others through strict conformity, while the secure position of socially accepted persons makes them less responsive to group pressures. The central concepts are social acceptance, rather than other aspects of informal status, and resistance to group pressure,

rather than deviate psychological tendencies or deviation from highly sanctioned mores, just as in Kelley's work. But whereas he has investigated the implications of high acceptance for resistance to group pressures only where such resistance contributes to the achievement of common goals, the prediction is here extended to include any resistance to group pressures, as long as it does not directly conflict with the achievement of important collective goals.

The rest of the paper is more exploratory. There is a multitude of attitudes, judgments, and practices for which group standards and deviations from them can be established. Patterns of deviation from a number of group standards are investigated; these fall into two broad categories, the orientation of caseworkers toward the clients who receive public assistance, and their orientation toward their supervisor. Since the findings support the hypothesis in some cases but not in others, an interpretation is suggested that accounts for these differences by specifying the type of orientation which makes unintegrated workers most subject to group pressures. Finally, the question is explored of how a person's orientation to his clients and work is related to his tendency to differ from the majority of his work group in his orientation to his supervisor.

PROCEDURE

The data analyzed in this paper are part of a study of a public assistance agency in a large American city.¹ The caseworkers had the duty to determine whether new applicants who came to the agency were eligible for public assistance, and whether old recipients continued to be eligible. This involved visiting the clients in their homes and a considerable amount of paper work in the office. The caseworker was also expected to see that his clients obtained special services they needed, such as medical, legal, or employment service, and to help and encourage clients to become self-supporting again. Caseloads were heavy—on the average, over 120 cases per worker—which limited opportunity for providing casework service. Most workers had only a college degree and no professional social work training. Nearly half the workers were women, and a third of them were Negroes. Each work group of five or six workers had a supervisor.

After a period of observation, when data were collected which are not used here, most of the members of twelve work groups—60 caseworkers—were interviewed either in a private room at the agency or in their homes.

¹ I am indebted to the Social Science Research Committee, University of Chicago, for providing funds for this pilot study, and to Philip M. Marcus for assisting in the collection and analysis of the data.

The interview, administered by the author and one assistant, contained both precoded and open-ended questions. The aspects of the caseworker's orientation selected for the deviation analysis are those that seem to be most significant in the context of a bureaucratic organization devoted to public service, such as detachment versus involvement, attitude toward clients, service versus procedure orientation, loyalty to the supervisor, and evaluation of the supervisor's approach and work. Negative as well as positive results of the analysis are reported.

In the study of the relationship between social status and deviation, care must be taken not to confuse a correlation between social status and certain opinions with a correlation between social status and a tendency to hold *deviant* opinions. An illustration will make this clear. In New York City, where most people are Democrats, one would find a disproportionate number of middle-class persons who are Republicans, that is, who hold deviant political opinions. But to conclude from this that high socio-economic status is associated with a tendency toward deviancy in political attitudes would not be justified, as the data from a small community in up-state New York would demonstrate. There, where most people are Republicans, low socio-economic status would be associated with deviant political attitudes, that is, with Democratic attitudes. These data show that low socio-economic status encourages a favorable attitude toward the Democratic Party, but they reveal no relationship between socio-economic status and political deviancy. But it is only if we examine communities with contrasting climates of opinion that we can distinguish between the correlates of certain opinions and the correlates of deviancy.

The hypothesis that an integrated status among peers promotes resistance to group pressure implies that integration is related to deviancy as such, regardless of the specific nature of the group climate. As long as one examines only conformity with a given value, say, a positive orientation to clients, one can never be sure whether the correlations found indicate who the *deviants* are or who the persons are who tend to hold negative attitudes toward clients. To test the hypothesis, therefore, requires analysis of the relationship between integration and deviancy in *contrasting group climates*.

For this purpose, the following procedure is used: On any item under consideration, the twelve work groups are first divided into those where the majority expresses one opinion and those where the majority expresses the opposite opinion. This constitutes the measure of group climate, and the assumption is that group pressures operate to produce conformity with the respective climate. Within each type of group, individuals are then divided on the basis of the opinion they have expressed; that is, individuals are

divided on the basis of the same empirical measure as that used for dividing groups. This furnishes four columns (see any table): the A's in groups where the majority are A's, the B's in A-groups, the A's in B-groups and the B's in B-groups. Finally, a third variable is introduced (in the rows) to ascertain whether it differentiates the deviants in both types of groups. For example, are integrated workers more apt than others to be B's in A-groups and also to be A's in B-groups?

A number of the varied connotations of the term *deviation* are deliberately excluded from consideration by this operational definition in terms of sheer difference from the group majority in respect to a single attitude or practice. It is intended merely as a sensitive indication of resistance to group pressure. Clearly, this measure does not identify deviant personality types, such as, for example, neurotic people. Neither does it refer to the deviant social roles of individuals who are defined by other group members as wrongdoers or outcasts. (To hypothesize that the most accepted group members tend to be deviants in the sense of outcasts would obviously be absurd.) It does not even show that a person may have violated an explicit taboo of the group. It only indicates that one aspect of the orientation or conduct of an individual fails to conform to that which prevails in his group. But if this occurs in an area that is at all salient, the person whose behavior differs will experience some pressure to conform, and his failure to do so is indicative of resistance to group pressure. On the other hand, if the item selected is not at all salient, differences are likely to be random and not to reveal a pattern of association with social characteristics. There is, however, the possibility that other factors than resistance to group pressure, such as ignorance, produce opinions that differ from that of the majority, as we shall see.

Social integration is measured sociometrically. Some workers called one another by their first names, but this was not the standard practice in the public assistance agency. Since it was not a mere convention, the use of first names implied social acceptance. If some members of his own work group reported being on a first-name basis with a person, he is defined as integrated, and if none did, he is defined as unintegrated. (Ideally, this index should be refined by taking into account the influence of background characteristics on first-naming, but analysis revealed these patterns to be too complex to make this feasible.) A second measure of social acceptance is the worker's popularity, as indicated by the fact that he was named by two or more respondents in the entire sample as a friend. Popularity and integration were significantly but not very highly related. The purpose of introducing the second measure is not to investigate the subtle distinction between integration and popularity, but to strengthen the evidence by using two different indicators of social acceptance.

ORIENTATION TO CLIENTS

An important element in the orientation of caseworkers in a complex organization is whether they remain detached toward their cases or become quite involved in their work. Respondents were asked how often they worry about their cases after working hours. Work groups where the majority answered "sometimes" or "often" were classified as highly involved, and those where the majority answered "rarely" or "never," as little involved or detached. The same criterion was used to divide individual workers within each type of group into those with high and low involvement. The data presented in Table 1 indicate that integrated workers deviated from the prevailing group climate in disproportionate numbers regardless of the nature of this climate.² They were slightly *less* apt to be involved (50 per cent) than unintegrated workers (62 per cent) in groups where involvement was high, and they were considerably *more* apt to be involved (50 per cent) than unintegrated workers (0 per cent) in groups where involvement was low. As a matter of fact, the integrated workers were not at all influenced by the degree of involvement of their co-workers; whether they were in groups where the majority was involved or where the majority was detached, half of them were involved. If the proportion of deviants in all twelve groups is computed, whether being a deviant means being involved or the the opposite, exactly one-half of the 34 integrated workers but only about one-fifth (5) of the 26 unintegrated ones deviated from the prevailing group climate, a difference that is significant on the .05 level.³

Popularity, as well as integration in the in-group, was associated with tendencies toward deviation. Popular workers were disproportionately often involved in groups where involvement was low and disproportionately rarely involved in groups where involvement was high. The same tendencies toward deviation of integrated and popular workers can be observed in respect to other aspects of orientation to clients and work. Whether a worker favored raising the assistance allowance for all clients or not, for example, proved to be a good measure of his attitude toward clients in general. Popular workers tended to deviate from the attitude toward clients that prevailed in their work group in disproportionate numbers, whether this attitude was positive or negative. Two-fifths of the 30 popular workers, in contrast to one-fifth of the 30 unpopular ones, expressed attitudes that departed from

² I have presented this table also elsewhere, as part of a methodological discussion (4).

³ The simplest way to test the significance of these differences is to combine the two columns of deviants and the two columns of conformists and use the chi-square test for a 2×2 table. An alternative procedure is to test the second-order interaction in the complete table (16, pp. 202-03). Since the results of the two tests did not appreciably differ, the simpler one (two-tailed) is used.

TABLE 1

Social Integration and Deviation in Involvement

	Group's Dominant Climate							
	High Involvement				Low Involvement			
	Individual's Involvement				Individual's Involvement			
	High %	Low %	Total %	N	High %	Low %	Total %	N
Position in work group								
Integrated	50	50	100	18	50	0	100	16
Unintegrated	62	38	100	13	0	100	100	13

The difference between rows in the total proportion of deviants is significant on the .05 level.

those of the group majority. Although these other relationships do not reach statistical significance (on the .05 level), seven out of eight manifest the pattern predicted by the hypothesis.

The cross-sectional data do not furnish direct evidence on the direction of influence that accounts for these relationships between acceptance among peers and resistance to group pressure. But it seems not very plausible that deviant tendencies are rewarded by greater acceptance and liking. The more probable inference, therefore, is that an integrated position increases the tendency to deviate from the prevailing group climate. Social acceptance by peers makes the work situation much more pleasant, provides support that reduces anxieties engendered by conflicts with clients and possibly with the supervisor, and facilitates obtaining advice when necessary. The unintegrated worker, consequently, has strong incentives to make himself more acceptable to others and avoid anything that may antagonize them, and this makes him highly responsive to group pressures. The worker whose integrated position gives him a feeling of security, on the other hand, is less apt to experience being restrained by the group and thus more apt to voice his opinion even if it differs from that of the majority. This interpretation implies that feelings of security and self-confidence are directly related to the tendency to deviate. Indeed, this seems to be the case. One-half of the 18 workers whose self-confidence is indicated by their conviction that they could work without supervision deviated in their involvement from the prevailing group climate, in contrast to less than one-third of the 42 less self-confident workers.

ORIENTATION TO THE SUPERVISOR

Caseworkers were asked to which supervisor they would want to be assigned if there were a reorganization and they could choose any supervisor.

Whether a respondent chose his present supervisor or not can be considered an index of loyalty to the supervisor. If groups are divided on the basis of whether the majority preferred to remain with their present supervisor or not, and individuals are divided in terms of the same criterion, it is again evident that integrated workers deviated from the majority in their group disproportionately often (Table 2). They were less inclined than unintegrated workers to be loyal to supervisors who commanded the loyalty of the majority of their subordinates, and they were more inclined to be loyal to supervisors who only commanded the loyalty of a minority. The proportion of deviants in all twelve groups is only 15 per cent among the 26 unintegrated

TABLE 2
Social Integration and Deviation in Loyalty to the Supervisor

	Group Majority's Attitude to the Supervisor							
	Loyal				Not Loyal			
	Individual's Attitude			N	Individual's Attitude			N
	Loyal %	Not %	Total %		Loyal %	Not %	Total %	
Position in work group								
Integrated	73	27	100	11	48	52	100	23
Unintegrated	100	0	100	13	31	69	100	13

The difference between rows in the total proportion of deviants is significant on the .05 level.

workers, but 41 per cent among the 34 integrated ones, which is a significant difference. Popular workers, too, were more prone to deviate from the group majority in their loyalty to the supervisor than unpopular ones.

Lack of integration among peers discourages deviations from the group majority in respect to loyalty to the supervisor as well as in respect to orientation to clients and work. The constraining influence of the work group appears to be less effective for those individuals who are fully accepted by their co-workers than for those who are not. To be sure, this may not be the case in laboratory groups, which are usually of such short duration that remaining an isolate is not too threatening. But a person spends so much time in his work group that remaining an isolate there is too punishing an experience not to be avoided if at all possible. This gives the unintegrated worker strong incentives for strict conformity, incentives that the integrated worker no longer has, since maintaining one's informal status is less demanding than achieving it. (These considerations seem to imply that newcomers would conform more strictly to group norms than workers who have been with the agency for a long time. Actually, however, seniority was not associated with deviant tendencies in regard to the orientations discussed so far.)

Another crucial aspect of a person's orientation to his supervisor is his evaluation of whether the supervisor is authoritarian or not. The descriptions workers gave of their own supervisor's approach to workers are classified as non-authoritarian—friendly, informal, helpful, accessible, willing to delegate responsibility—and authoritarian—unfriendly, domineering, interfering, overprotective, rigid about procedures. Work groups are divided on the basis of whether the supervisor was considered more or less authoritarian, and individuals are divided on the basis of whether they mention only non-authoritarian qualities of their supervisor or also authoritarian ones. There was an inverse relationship between the fact that a supervisor was generally looked upon as authoritarian and the loyalty of his subordinates. Despite this relationship, the patterns of deviation in respect to these two facets of the orientation to the supervisor were quite dissimilar.

Social acceptance among peers was not related to the tendency to differ from the group majority in judging the supervisor's approach to workers, regardless of which measure of acceptance is used. In all work groups, 31 per cent of the 26 unintegrated workers and 38 per cent of the 34 integrated ones deviated from the group majority in their evaluation of the supervisor's approach. Similarly, the proportion of deviants among the 30 unpopular workers (37 per cent) hardly differed from that among the 30 popular workers (33 per cent).

This negative finding raises the question of whether other dimensions of informal status, such as a worker's status as consultant among colleagues, are associated with deviant evaluations of the supervisor. Indeed, such an association can be observed, but it is neither significant nor in the expected direction. High-status consultants were slightly *less* apt to deviate from the group majority in judging the supervisor's approach than workers who were rarely consulted by colleagues. Further analysis indicates that these deviant tendencies are probably not due to the *social* status of consultant but to other correlated factors, namely, experience on the job and knowledge of it.

Seniority, which indicates accumulated experience in the agency, was related to deviancy in judging the supervisor's approach to workers (Table 3). Workers with less than three years of seniority were less apt than old-timers to express positive evaluations of supervisors whom the majority evaluated positively, but they were more apt than old-timers to express positive evaluations of supervisors who were generally criticized as authoritarian. In short, workers with few years of experience differed from the group majority disproportionately often in judging the supervisor's approach, regardless of whether the supervisor was authoritarian or non-authoritarian. Of the 18 workers with over three years' seniority, only 11 per cent made deviant judgments; of the 42 workers with less than three years' seniority, fully 38 per

cent did (significant on the .05 level). A worker's competence, as indicated by his supervisor's rating, had the same, but less pronounced, implications for deviation. Whether the supervisor was generally considered authoritarian or not, the more competent workers were more prone to be in accord with the majority judgment than the less competent ones.

In judging the supervisor's approach, experience on the job fostered deviant responses, but social acceptance did not. In other respects, however, acceptance was associated with deviant tendencies and experience was not. How can these findings be explained? Whether an individual likes his own supervisor or prefers another one is clearly a value judgment, and his orientation to his clients and work also rests on value judgments. When a worker de-

TABLE 3
Seniority and Deviation in Judging Supervisor's Approach

	Group Majority's Judgment of the Supervisor							
	Non-Authoritarian (+)				Authoritarian (-)			
	Individual's Judgment				Individual's Judgment			
	+	-	Total	N	+	-	Total	N
Seniority	%	%	%		%	%	%	
Under three years	50	50	100	20	27	73	100	22
Over three years	82	18	100	11	0	100	100	7

The difference between rows in the total proportion of deviants is significant on the .05 level.

scribes his supervisor's approach to subordinates, however, he makes a factual judgment; he does not express a preference but, presumably, reports how the supervisor actually behaves. Of course, the analytical distinction between value and factual judgment (15, pp. 45-60) is not reflected in pure form in everyday life. Factual considerations underlie value judgments—for example, why a person prefers a certain supervisor—and values influence factual judgments—for example, which supervisory practices a person defines as interfering rather than helpful. Nevertheless, judgements that can be potentially validated in terms of an objective reality differ from those that cannot be. Social pressures will exert a more pronounced influence in the latter case than in the former, because the climate of opinion in which an individual lives constitutes, in effect, the social reality that validates his value judgments (6).

Factual judgements are more or less correct. Experience and knowledge, therefore, have much bearing on them. Caseworkers with many years of experience in the agency, which often has involved working under different

supervisors, are better qualified to make accurate judgements of the supervisor's approach than more recent arrivals, who may mistake a gruff manner for authoritarianism or a manipulative approach for genuine friendliness. Insufficient experience prevents a worker from being discriminating in judging the supervisor's approach, and lack of discrimination in describing different styles of supervision manifests itself in a pattern of deviant responses. Workers with less than three years' seniority did not describe non-authoritarian and authoritarian supervisors very differently, as Table 3 shows; half of them evaluated non-authoritarian supervisors in positive terms, and 30 per cent evaluated authoritarian supervisors in positive terms. Old-timers discriminated more sharply; 82 per cent evaluated non-authoritarian and none of them evaluated authoritarian supervisors in positive terms. Since these judgements are largely governed by the actual conduct of the supervisor, on the one hand, and the ability to make accurate judgements of this objective reality, on the other, the influence of group pressure is minimized. But value orientations are not so much affected by experience or knowledge as by social forces, and how strong the influence of the group climate on the individual's orientation is depends on whether his position in the group is sufficiently secure to permit him to resist group pressures.

Social factors, however, also may play a part in producing the difference found between old-timers and newcomers. A subordinate's evaluation of his superior is confidential information, which is not freely revealed for fear that some criticisms may get back to the superior. Old-timers who have known one another for many years may well discuss their supervisor, and thus reach consensus about him, but they are not likely to share their judgement with newcomers, whom they cannot yet trust. This reluctance to communicate to the newcomer the group's exact judgment of the supervisor combines with his lack of experience to create his greater tendency to deviate in his judgement from the presumably correct one of the majority. No such barriers to communication exist as far as orientation to clients is concerned. During the first few months, newcomers sometimes went into the field with old-timers and often discussed their problems with them, and the latter had no reason to conceal from the former what they considered to be the proper attitudes to clients.

IMPLICATIONS OF ORIENTATION TO CLIENTS AND WORK

Experience and social acceptance were not the only factors that had some bearing on patterns of deviation in the orientation to the supervisor; a worker's orientation to his clients and work did too. Generally, positive attitudes toward clients and service seemed to be associated with resistance to group pressure in the orientation to the supervisor. For example, workers

who unconditionally favored raising the assistance for clients differed in disproportionate numbers from the group majority in their overall attitude (loyalty) to the supervisor, and also in their judgment of his efficiency. Similarly, deviant judgements of the supervisor's efficiency were somewhat more prevalent among workers whose descriptions of a day in the field reveals them to be oriented toward casework service, than among those primarily concerned with checking eligibility procedures, as Table 4 shows. If the supervisor was considered efficient by the majority, service-oriented workers questioned his efficiency disproportionately often, and if he was generally not considered particularly efficient, service-oriented workers described him as

TABLE 4
Orientations to Clients and Deviation in Judging Supervisor's Efficiency

	Group Majority's Judgment of the Supervisor							
	Efficient				Not Efficient			
	Individual's Judgment				Individual's Judgment			
	Eff. %	Not %	Total %	N	Eff. %	Not %	Total %	N
Orientation to clients								
Procedure	75	25	100	12	6	94	100	16
Service	50	50	100	8	29	71	100	24

The difference between rows in the total proportion of deviants is significant on the .10 level.

efficient disproportionately often. Nearly one-third (11 of 32) of the service-oriented workers, in contrast to only one-seventh (4 of 28) of the procedure-oriented ones, deviated in making these judgments from the majority of their respective work groups. Although these relationships are not significant on the .05 level, the fact that nine of twelve examined assume the predicted pattern is at least suggestive.

A positive orientation toward clients and casework service supplies a worker with two reference groups—his clients and professional social work. It makes the reactions of his clients more significant for the gratification he obtains from his job, and it makes it more important for him to live up to professional casework standards. To be sure, the demands made on him by these two reference groups sometimes conflict; the one, for example, promotes involvement with clients, whereas the other discourages it. Nevertheless, the existence of two alternative frames of reference makes a worker less responsive to the pressures of his work group and induces him to discriminate along different lines from those conventionally employed by subordinates in evaluating the supervisor. A supervisor who is otherwise liked might have such negative attitudes toward welfare clients that he alienates service-

oriented workers, and one whose approach to subordinates is objectionable might know so much about casework that service-oriented workers appreciate him. Since external reference groups make workers oriented to casework service somewhat independent of the immediate work group and furnish them with distinctive criteria for judging supervision, they disproportionately often deviate from the majority in evaluating their supervisor.

The descriptions of the supervisor's approach to workers, however, do not reveal such a difference in the tendency to deviate between workers with contrasting orientations to clients. As a matter of fact, one finding seems to imply that service-oriented workers were *less* prone than others to deviate in this respect, in contrast to their *greater* tendency to deviate in other respects. Workers who were willing to delegate certain responsibilities for buying clothing to their clients were less apt than workers who were unwilling to do so to deviate from the group majority in judging the supervisor's approach, regardless of whether his approach was generally considered to be authoritarian or the opposite. At first, this was interpreted to show that procedure-oriented workers—that is, those who want to continue to keep a close check on the spending of clients—are more likely than service-oriented ones to deviate. But an alternative interpretation of the finding is that workers who are greatly interested in exercising responsibilities—and who, therefore, do not want to give up any responsibilities they now have—are particularly prone to deviate. To test this second inference, it is necessary to determine whether more direct measures of attitude to responsibility are similarly related to deviation in describing the supervisor's approach to workers.

When a caseworker initially determined the amount of assistance an applicant would receive, his decision had to be approved by the supervisor. Respondents were asked whether they would want to make these decisions without special supervisory approval. Workers who were definitely interested in assuming these additional responsibilities deviated from the group judgement of the supervisor's approach to workers in disproportionate numbers, as Table 5 shows. They were less apt than others to evaluate non-authoritarian supervisors in positive terms but more apt to evaluate authoritarian supervisors in positive terms. In sharp contrast to other workers, those positively oriented toward responsibility were just as inclined to mention positive things about the approach of an authoritarian supervisor (43 per cent) as about that of a non-authoritarian one (42 per cent). Taking the twelve groups together, 53 per cent of the 19 responsibility-oriented workers and 14 per cent of the 28 other workers made judgements that deviated from the group majority, a highly significant difference. Ambitious workers (those who expected a promotion), just as responsibility-oriented ones, were more likely than others to deviate from the majority of their work group in describing

the supervisor's approach to workers, although the differences in this case were considerably less pronounced.

These findings suggest that a concern with assuming responsibilities and becoming promoted interferes with a person's ability to discriminate between different supervisory practices. A position of superior authority assumes such significance in the eyes of the ambitious individual that he becomes anxious lest he appear irresponsible when evaluating it. Such a worker may approve of the supervisor's approach but, wary not to sound naive and uncritical, qualify his approval so much that it turns into a criticism. Or he may con-

TABLE 5

Orientation to Responsibility and Deviation in Judging Supervisor's Approach

	Group Majority's Judgment of the Supervisor							
	Non-Authoritarian (+)				Authoritarian (-)			
	Individual's Judgment				Individual's Judgment			
	+	-	Total	N	+	-	Total	N
	%	%	%		%	%	%	
Orientation to responsibility *								
Positive	42	58	100	12	43	57	100	7
Negative	75	25	100	12	6	94	100	16

The difference between rows in the total proportion of deviants is significant on the .01 level.

* Not all workers were asked this question. Hence, the total N is less than 60.

sider the supervisor authoritarian but be so anxious to prove that he understands the problems the supervisor faces that he gives a fundamentally positive description of his approach. The anxious concern of responsibility-oriented workers to make careful judgements of supervisory practices seems to have, paradoxically, the result that their judgements, at least in the aggregate, fail to discriminate between quite distinct styles of supervision.

The findings also permit, however, another interpretation, which centers neither upon deviancy nor upon lack of discrimination, but upon the relationship between the supervisor's approach and that of the workers who describe it. Workers who enjoy exercising responsibilities and are reluctant to delegate them to clients are somewhat akin in their approach to authoritarian supervisors, while workers who have little interest in increasing their power of discretion and who are willing to delegate responsibilities have an orientation that is more similar to that of non-authoritarian supervisors. The fact that a worker's approach is similar to that of his supervisor apparently introduces a favorable bias into his description of supervisory practices. The data presented in Table 5 show that supervisors were evaluated in more favorable terms by workers whose approach was similar to their own than by workers

with a different approach; that is, non-authoritarian supervisors were described in positive terms most often by workers not interested in responsibility (75 in contrast to 42 per cent), while authoritarian supervisors were most often so described by responsibility-oriented workers (43 in contrast to 6 per cent).

This interpretation implies that authoritarian supervisors would also be more likely to command the loyalty of responsibility-oriented workers than that of others, while non-authoritarian supervisors would be less likely to command the loyalty of the former than that of the latter. Indeed, the data reveal the predicted relationships, although differences are too small to be statistically significant. Whichever measure of responsibility orientation is used, workers oriented toward responsibility are more likely than other workers to be loyal to authoritarian supervisors, but less likely than others to be loyal to non-authoritarian supervisors.

CONCLUSIONS

The first problem in the study of deviation from group standards is to distinguish factors that actually have implications for deviation as such from those that have implications for identification with certain values which happen to be infrequent in the groups under consideration. For this purpose, a procedure is used that indicates whether a variable is associated with deviant tendencies *both* in groups where one climate of opinion prevails and in groups where the opposite climate of opinion prevails.

Patterns of deviation in a group are governed by forces external to it as well as by internal forces. Nonconformity with the prevailing values in a group can often be considered conformity with the standards of another reference group, as Merton, freely acknowledging his debt to Cooley, has pointed out (12, pp. 357-59). Clients and the profession of social work constituted reference groups for those caseworkers in a public assistance agency who were interested in service to clients, but not so much for those primarily concerned with eligibility procedures. Superiors in the organization probably were reference groups for workers oriented toward exercising responsibility and upward mobility, but not for workers without such ambitious orientations. The influence of all these reference groups was reflected in the greater tendency of workers who were oriented to them, than of those who were not, to deviate from the majority in their judgments.

There is apparently an important difference between conformity to the factual judgements of the group majority and conformity to its value judgements. Even though the factual judgements that have been examined—the workers' descriptions of the supervisor's approach—are by no means free

from value elements, the fact that there is an objective reality—the supervisor's behavior—in terms of which these judgements can be evaluated as more or less correct is of great significance. It has the result that the ability to discriminate and the criteria used in discriminating have a crucial bearing on the tendency to conform to the majority's factual judgements. Workers with less than three years of experience were less capable of discriminating between various styles of supervision than experienced old-timers, probably not only because they did not have the required knowledge, but also because old-timers did not trust newcomers enough to communicate to them their opinion of the supervisor. Workers with low seniority, consequently, expressed deviant judgements of the supervisor's approach disproportionately often. Responsibility-oriented workers, too, tended to make less discriminating and thus more deviant judgments of the supervisor's approach than others. It is interesting that lack of experience and a responsibility orientation had similar effects on these judgements, particularly since inexperienced workers were slightly less apt than others to be oriented toward responsibility. Perhaps it is the difference between a person's level of aspirations and his actual abilities that interferes with making discriminating judgements. A responsibility-oriented worker might demand of himself judgements of such high quality that his ability to live up to his own standards falls as far short as the ability of the inexperienced worker to live up to his considerably lower standards. Such inability to live up to one's own expectations may well engender anxieties that disturb an individual's powers of discrimination.

Value orientations, which do not rest on factual judgments, are most subject to social pressures. A worker's tendency to deviate from the other members of his work group in his orientation to clients was not influenced by his experience, but it was related to his position in the group, and so was his tendency to deviate in his loyalty to the supervisor. Workers who were integrated in their peer group were more apt to express deviant value orientations than unintegrated workers, as had been hypothesized. A person who is not fully accepted in his work group has strong incentives to become better integrated and prove himself acceptable by conforming to the prevailing values of the rest of the group. A secure position among colleagues relieves an individual of this concern with improving his interpersonal relations on the job and thus frees him somewhat from group pressures. It permits him to let his thinking and acting be guided by his own principles, which he has brought to the group from previous experience, and which often are the standards of his reference groups, even if this makes him in some respects a deviant. All persons have many reference groups, the standards of some of which are apt to conflict with those of their work group. How much influence his other reference groups exert on an individual depends, in large part, on how

secure his position in the work group is and how well he can therefore resist group pressures.

The conclusion that *high* acceptance in a group is associated with deviation from group standards must be qualified in view of the findings of some other studies that *low* acceptance is associated with deviancy. First, purely individualistic kinds of deviancy, such as neurotic behavior, are probably fostered by social isolation and not by integration (8). Second, the relationship may be curvilinear, and both the most and the least accepted members of groups may deviate more often than those in intermediate positions (5), but the few cases in the sample made it impossible to investigate this possibility. Third, the rejected person in a neighborhood can cease to have social contacts with his neighbors and seek his companions elsewhere, thus removing himself from the control of the neighborhood group. But the rejected person in a work group is required by his job to continue to interact with co-workers, so that the only practical solution to his problem is to try to improve his position through strict conformity. This would explain why the unintegrated members of neighborhoods were found to deviate from group norms disproportionately often (7), whereas the unintegrated members of the work groups in the public assistance agency deviated disproportionately rarely. Finally, social norms that pertain to the collective goals of the work group, such as restriction of output, are too significant for the group to permit anybody to violate them. Hence, only the outcasts are apt to do so; that is, the unintegrated isolates are most likely to deviate from these salient norms (9, pp. 140-44; 2, pp. 158-59, 166-67). But the prevailing climate of opinion in a work group, as distinct from its most salient values, exerts less restraints upon the integrated member, who has proven his acceptability and can be permitted some freedom to deviate, than upon the unintegrated one, who has yet to prove it.

Manuscript received: July 30, 1959

Revised Manuscript received: October 26, 1959

Peter M. Blau

Department of Sociology

University of Chicago

Chicago 37, Illinois

REFERENCES

1. Bales, R. F. and P. E. Slater, "Role Differentiation in Small, Decision Making Groups," in T. Parsons and R. F. Bales, *Family, Socialization and Interaction Process*, Glencoe: Free Press, 1955.
2. Blau, P. M., *The Dynamics of Bureaucracy*, Chicago: University of Chicago Press, 1955.
3. Blau, P. M., "Social Integration, Social Rank, and Processes of Interaction," *Human Organization*, 1959-60, 18, 152-157.

4. Blau, P. M., "Structural Effects," *American Sociological Review*, 1960, 25, 178-193.
5. Dittes, J. E., and H. H. Kelley, "Effects of Different Conditions of Acceptance upon Conformity to Group Norms," *Journal of Abnormal and Social Psychology*, 1956, 53, 100-107.
6. Festinger, L., "A Theory of Social Comparison Processes," *Human Relations*, 1954, 7, 117-140.
7. Festinger, L., et al., *Social Pressures in Informal Groups*, New York: Harper, 1950.
8. French, R. L., "Sociometric Status and Individual Adjustment among Naval Recruits," *Journal of Abnormal and Social Psychology*, 1951, 46, 64-72.
9. Homans, G. C., *The Human Group*, New York: Harcourt, Brace, 1950.
10. Kelley, H. H., and M. M. Shapiro, "An Experiment on Conformity to Group Norms Where Conformity is Detrimental to Group Achievement," *American Sociological Review*, 1954, 19, 667-677.
11. Merei, F., "Group Leadership and Institutionalization," *Human Relations*, 1949, 2, 23-39.
12. Merton, R. K., *Social Theory and Social Structure*, rev. ed., Glencoe: Free Press, 1957.
13. Newcomb, T. M., "Attitude Development as a Function of Reference Groups," in E. E. Maccoby et al., *Readings in Social Psychology*, 3rd ed., New York: Holt, 1958.
14. Riecken, H. W. and G. C. Homans, "Psychological Aspects of Social Structure," in G. Lindzey (ed.) *Handbook of Social Psychology*, Cambridge: Addison-Wesley, 1954, Vol. II.
15. Simon, H. A., *Administrative Behavior*, New York: Macmillan, 1945.
16. Snedecor, G. W., *Statistical Methods*, 4th ed., Ames: Iowa State College Press, 1946.

Psychological Dimensionality and the Distribution of Rank Order Agreement Among Judges

WILLIAM L. HAYS, *University of Michigan*

A problem of fairly frequent occurrence in social-psychological research concerns the similarity in preferences expressed by two individuals or judges. Often the measurement of such preference takes the form of a ranking by each judge of a set of objects or statements in terms of "desire to own" or "value," a set of attitude items in terms of degree of endorsement, a rank-order of sociometric preferences among peers, etc. The measure of the similarity of two judges is then taken as rank correlation between the judges in terms of Spearman's rho or Kendall's tau. In other instances, the agreement among a set of m judges may be obtained and a measure of concordance, such as Kendall's W , calculated, which is then interpreted as a reflection of the degree of homogeneity of preference existing among the judges. A representative study of this kind is that reported by Newcomb (7), in which each subject rank-ordered the six Spranger "values" in terms of relative importance to himself. The subjects then also estimated the rank order of importance to each of a group of peers. From these rankings, rho coefficients were computed, expressing the degree of both estimated and actual agreement.

The consideration to be raised here is important when the experimenter wishes to go further and make a statement about the statistical significance of the rank correlation or concordance measures. It will be shown that the use of the usual significance tests for such measures necessitates an assumption about the psychological structure which the judges impose upon the objects of judgement; that is, an assumption is made about the psychological dimensionality of the particular set of stimulus objects ranked. Before this may be shown, however, it will be convenient to introduce some terminological distinctions.

The general data-model to be adopted here is that of Coombs (2, 3). Coombs starts with the postulate that any object of judgment for an individual may be thought of as a point (or more correctly, as a vector) in a psychological "stimulus space" of some dimensionality r . That is, the object has certain relevant psychological attributes for the judge. The position of object k , for an individual i , on attribute j may be symbolized by x_{ijk} .

Now Coombs (3) further distinguishes between two kinds of "tasks" which the judge performs; he calls these Task A and Task B. Task A asks the judge to evaluate the stimuli relative to some "ideal stimulus" which he would most prefer or endorse; the task of the judge is to describe his preferences

for the stimuli, in terms of their similarity to his ideal. Task B, on the other hand, asks the judge to describe the stimulus objects in terms of each other: Task B asks the judge to tell the relative similarities of the stimulus objects to each other. Unlike Task A, in Task B the "ideal stimulus" of the judge is completely irrelevant. Any responses to attitude, value, or similar personal preference stimuli thus qualify as Task A judgments in which the subject is expressing his *own* psychological "position" (the position of his own "ideal"), while Task B judgments are best typified by responses to stimuli in psychophysical "methods of similarity."

In somewhat more formal terms, the ideal stimulus of judge i may be designated as c_i , and it too has positions on the relevant psychological attributes of the stimulus space; thus, c_{ij} is the status or "loading" of the ideal of i on the attribute j . Tasks A and B may be characterized by the following observational equations, linking the phenotypic judgement behavior of the individual to the genotypic positions of the stimuli and ideal in the attribute space:

Task A: Given two stimuli k and l , k is preferred to l if and only if

$$\sum_j (c_{ij} - x_{ijl})^2 > \sum_j (c_{ij} - x_{ijk})^2.$$

This states that the judge prefers the stimulus which is closer to his ideal in terms of psychological distance in the stimulus space.

Task B: Given three stimuli k , l , and m , k and l are judged more similar than k and m if and only if

$$\sum_j (x_{ijk} - x_{ijm})^2 > \sum_j (x_{ijk} - x_{ijl})^2.$$

That is, the reported relative dissimilarity between pairs of stimuli is always proportional to the relative distances between stimuli in the stimuli space.

In short, there are two quite different kinds of information which judges may provide about a set of stimulus objects, *depending on whether we "ask" them about similarity of stimulus objects to their ideal object or about similarity of stimulus objects to each other*. When one asks for preference or preference-like judgments, Task A is required. Similarity and other scaling methods (such as the familiar Thurstone scaling procedures) ask for Task B judgments. Underlying most scaling methods is the assumption that meaningful sets of stimulus objects will yield relatively small dimensionality, and the further assumption that judges are able to view stimuli in much the same way, within certain prescribed limits of error. The usual scaling methods treat stimuli as though the dimensionality, relevant attributes, and the perceived psychological structure of stimulus objects will be at least roughly comparable from subject to subject; without such assumptions attempts to arrive at

psychological scales utilizing pools of judges would be pointless. Thus, for the moment, the fact that the relevant attributes and the psychological relationships among a set of objects are, ultimately, peculiar to judge i will be ignored. Rather it will be assumed that judges can show at least a substantial agreement about the relations among a set of stimuli; this would mean virtually complete agreement on Task B, and disagreement on Task A due only to judges' own "ideal" positions.

Now, for the main argument to be presented here, suppose that there are two judges, each of whom rank-orders his preferences for three objects of judgement, A, B, and C. Suppose further that for our two judges, the three stimuli have only *one* relevant psychological attribute (i.e., the dimensionality of the stimulus space is 1). Let us take the order of the stimuli on this attribute to be $A > B > C$; or $x_{iA} > x_{iB} > x_{iC}$, for any judge i . The order of the three stimuli on the attribute may be pictured as the order of three points on a line, as in Figure 1. When a judge gives his preference, to what

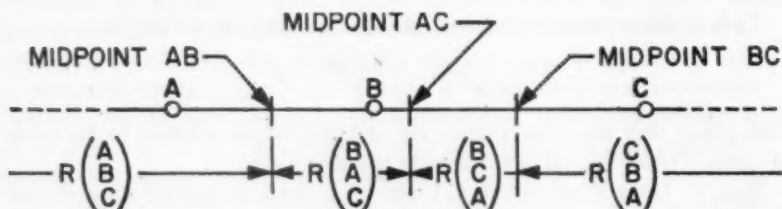


FIGURE 1

Possible Rank Order Regions Generated by Three Points in One Dimension

extent does the position of the stimuli on the single attribute delimit the preference orders he may have? The answer is, to a large extent. Figure 1 shows the possible number of regions into which the ideal stimulus for an individual c_i might fall, given three points in one dimension. There are only *four* such regions, each showing a different rank order of distances from any point in the region to the three stimulus points. Each region thus represents

a different order of preference. The symbol $R \begin{pmatrix} A \\ B \\ C \end{pmatrix}$ denotes that an individual with ideal stimulus located in that region must prefer A to B to C. On the other hand, $R \begin{pmatrix} B \\ C \\ A \end{pmatrix}$ is the region from which judgments of preference of B to C to A must emanate.

If the judges' ideal stimulus fell at one of the "midpoints" which separate the line into regions, he should prefer two of the stimuli equally, as the mid-

points are in a sense "regions of indifference," since they are equally distant from two stimuli. Note that the unidimensionality of the attribute space in this situation precludes the possibility that all $n!$, or six, permutations in order of preference may occur. In the example at hand, the order of preference $R \begin{pmatrix} A \\ C \\ B \end{pmatrix}$ and $R \begin{pmatrix} C \\ A \\ B \end{pmatrix}$ may not occur. Figure 2, for example, shows three

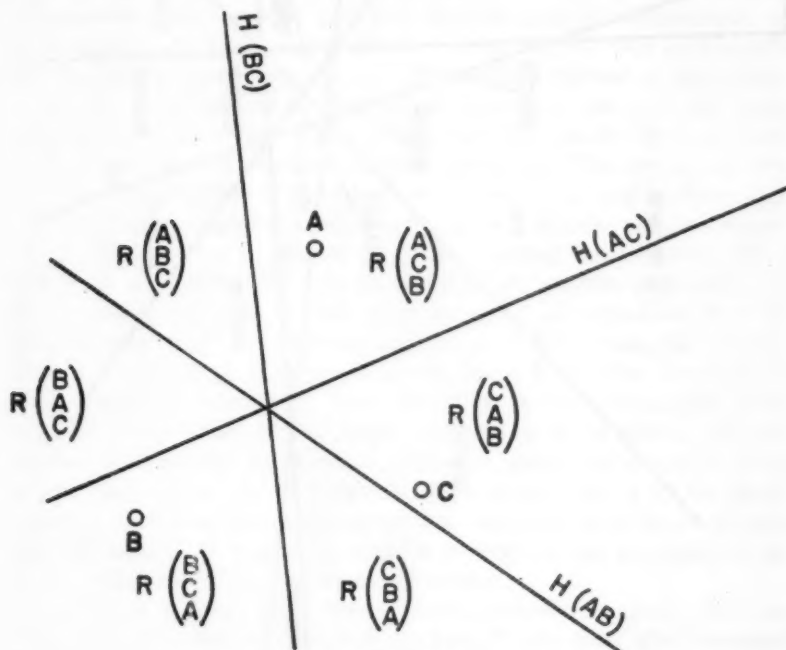


FIGURE 2

Possible Rank Order Regions for Three Points in Two Dimensions

stimulus points in two dimensions, and here all six ($n!$) of the orders of preference may occur, since the ideal stimulus of the individual may fall into any one of six possible rank order regions.

For a more complicated example, one may turn to Figure 3, which represents four stimulus points in two dimensions. In this instance, there are 18 different orders which the subject may show for this configuration in two dimensions. In neither one nor two dimensions for four stimuli, however, are all $n!$ (or 24) rank order permutations possible as orders of preference for a judge. It has been shown that, in general, all $n!$ different rank order permu-

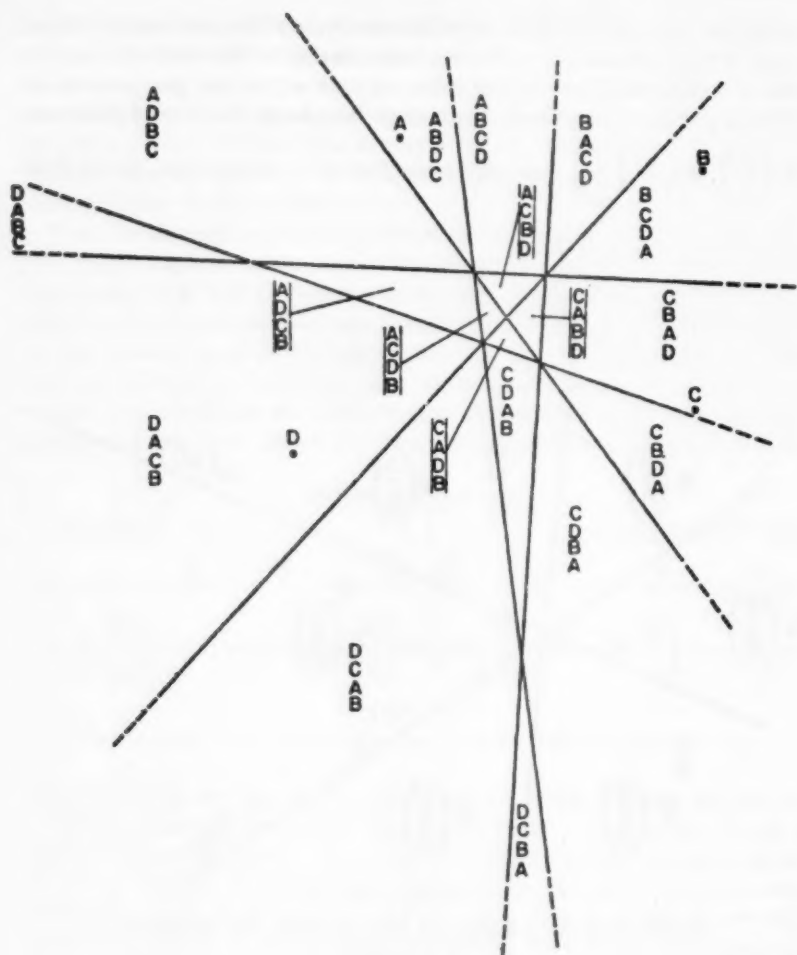


FIGURE 3

Rank Order Regions Generated by Four Points in Two Dimensions

tions of any n stimuli may occur *only* when the stimuli require at least $n-1$ dimensions (1, 5). For any dimensionality less than $n-1$, *some* rank orders become impossible. For example, all rank orders of three stimuli may occur only when the dimensionality is two, all orders of four may occur only when the dimensionality is three, all orders of five when the dimensionality is four, and so on.

Thus the number of rank orders of preference which may occur is bound up with the dimensionality of the psychological "space" of the stimuli. *The practical implication of this fact is that statistical tests of rank order agreement or concordance between a pair or among many pairs of judges may be seriously biased.* In asking for rank orders from judges, we usually require Task A, some form of order of "preference," and wish to see if judges' orders of preference agree; however, the probability distribution of various degrees of agreement under "random" conditions depends upon the dimensionality of the stimulus space. *The sampling distributions of measures such as Kendall's tau, average tau, Spearman's rho, and the Kendall coefficient of concordance (4, 6) are all worked out in terms of the assumption that each and every permutation in rank order is not only possible, but equally likely to occur in the population from which the rankings were drawn.* This can be true only when the dimensionality of the space is $n-1$. Moreover, it may be shown that $n-1$ dimensions are sufficient to describe any set of n stimulus objects whatever (1, 5); thus $n-1$ may be referred to as the "maximal dimensionality" for n objects. *In short, when used with Task A data, the sampling distributions of these common measures of rank order agreement all presuppose that the dimensionality of the stimuli is maximal (i.e., $n-1$).* This means that not only Task A but also Task B judgments of the judges in question should be, in effect, "randomly scattered"; there should be no true concordance either in terms of the ideals of the judges (the matter under study), but also minimal communality (i.e. minimal similarity) among the stimuli in terms of what they "mean" for the judges. In effect, in order to test for the significance of rank order preference agreements, one must make an assumption that the stimuli are maximally complex in terms of the psychological attributes which the judges view them as possessing.

In practice, however, one is almost always interested in stimuli which are rather simple in their psychological structure. We are most often interested in stimuli such as attitude items which "scale" or give other evidence of unidimensionality when studied by Task B methods. We rely on the fact that judges can agree on fairly restricted psychological meanings for, and relations among, stimuli. Hence, it seems that there is potentially a paradox here: In order to test for the statistical significance of Task A judgments, we assume that the psychological dimensionality of the objects judged is maximal; however, in choosing the objects to be judged, we often rely on those for which Task B judgments give evidence of minimal dimensionality in terms of scalability or communality. It is patently advantageous from an experimental point of view to have stimuli which seem to vary systematically along only one or a few attributes.

The effects of less than maximal dimensionality upon the sampling distribu-

tions of measures of rank agreement are quite easy to show. Consider for a moment the simple situation in which two judges give preference rankings for four objects. Table 1 shows the exact probabilities associated with various

TABLE 1

Probabilities for the Random Occurrence of Values of Kendall's Tau for $n = 4$ and Various Dimensionalities

τ	-1.00	-.67	-.33	0	.33	.67	1.00
1 Dimension	.041	.082	.122	.163	.204	.245	.143
2 Dimensions	.037	.111	.185	.241	.222	.148	.056
3 Dimensions (Kendall, 6)	.042	.125	.208	.250	.208	.125	.042

values of tau as reported by Kendall, under the assumption that each of the 24 rankings are equally likely (dimensionality equal to three). Along with these probabilities are shown the exact probabilities of the different taus when the dimensionality is 1, and each ranking is equi-probable and randomly selected. Note the overall effect is to skew the distribution to the left, making relatively high agreement have higher probability than low agreement. The expected value of tau for the unidimensional case is .205 rather than zero, the usual expectation in a test for significance of τ . Table 1 also shows the distribution of exact probabilities for the two-dimensional case (also with randomly selected equi-probable rankings), as shown in Figure 3. Note that for two dimensions the distribution is less severely skewed; the expected value is .056. *The extent of the bias toward agreement for a given dimensionality seems to be more severe the larger the number of objects.* This is illustrated in Table 2 by the probabilities of various values of tau for five stimuli in one dimension. Here the expected value of tau is .28.

What is true of tau is true of the Spearman rho as well. Table 3 contrasts the exact probabilities for rho for the one-dimensional and the three-dimensional case for four stimuli. In the one-dimensional case the expected value for rho turns out to be .266. Furthermore, the distribution of Kendall's W and of average tau as measures of concordance show the same sorts of trends: a more marked skewness to the left and a shift in expected value toward high agreement with lowered dimensionality.

For the sake of simplicity, this argument has been carried out as though the judges actually would agree about the psychological similarity among the stimulus objects, had they been given Task B to perform. It is not, however, necessary to limit the argument to this case. Even granting that the judges would disagree more or less about the rank order relations among the stimuli in Task B judgements, in any available rationale for combining judgments

TABLE 2

Probabilities for Kendall's Tau with $n = 5$ and Dimensionalities One and Four

τ	Dimensionality 4 (Kendall, 6)	Dimensionality 1
1.00	.008	.091
.80	.034	.166
.60	.075	.149
.40	.125	.133
.20	.166	.116
.00	.184	.099
-.20	.166	.083
-.40	.125	.066
-.60	.075	.050
-.80	.034	.033
-1.00	.008	.016

across judges to estimate the common psychological dimensionality, maximal disagreement among judges would still require and be required by a dimensionality of $n-1$ for the stimuli, such that each and every reported rank order by a judge is equally likely. Thus, if there is any consensus at all among the judges, and any judge tends to view the stimuli as having less than maximal dimensionality, the usual sampling distributions and tests of significance for rank order agreement will be in error.

All of the foregoing points toward the question of what, exactly, should a test of significance of preference agreement mean? One might say that it tests the hypothesis that each judge acts in a purely haphazard manner in giving his order of preference—that each judge takes his preference order “out of a hat” so to speak. If this is the interpretation to be given the significance test, then the objections raised here are invalid, since there is no implication, in

TABLE 3

Probabilities for Spearman's Rho with $n = 4$ and Dimensionalities One and Three

ρ	Dimensionality 3 (Kendall, 6)	Dimensionality 1
1.00	.042	.143
.80	.125	.245
.60	.041	.041
.40	.167	.163
.20	.083	.041
.00	.084	.082
-.20	.083	.041
-.40	.167	.082
-.60	.041	.041
-.80	.125	.082
-1.00	.042	.041

such a null hypothesis, that the judge pays any attention at all to the stimulus objects. If he is making up orders of preference out of thin air, then the assumption of equi-probability of the $n!$ possible permutations in rank is as good as any. It is difficult to see, however, why one would usually wish to entertain such a null hypothesis. The remarks made here are addressed to those who wish to consider that their judges do attend to the stimuli and that preference rankings do validly reflect the relation of the judge himself to those stimuli. The question to be asked of a significance test is then, "Are these judges more alike (or more different) than one has a right to expect individuals judging such objects to be?" When rephrased in this way, the importance of dimensionality as a factor in the expected agreement begins to emerge more clearly.

Perhaps the need is for ways to study the psychological dimensionality of stimulus spaces for groups of judges (Task B data) before questions of Task A agreement are taken up. Quite apart from any statistical considerations, such Task B studies might help to clarify the sources of interjudge agreement as "low stimulus dimensionality" or "nearness of ideal positions" or both. According to the argument given here, a high judge-pair Task A agreement should get more "weight" in a high-dimensional situation than in a low-dimensional one.

Furthermore, if these considerations are valid for the case of agreement between two judges, they should hold with even more force for the agreement with a judge's own preference and the preferences he estimates for others. Here it seems reasonable to suppose that a judge will attribute to another a stimulus space which is more or less consistent with his own view; we generally seem to operate on the assumption that others view the "objective" world in much the same fashion as we do. Unless the dimensionality of the judge's own space is maximal, the expected agreement between the ranking for himself and any estimated other *should* be greater than zero. Incidentally, it is rather striking that the distributions of the rho measures of estimated agreement reported by Newcomb (7) uniformly show the skewness to the left and the elevated means to be expected from stimulus spaces of less than maximal dimensionality. Newcomb shows that this tendency becomes more pronounced with general attraction toward, and familiarity with, the estimated other. Perhaps the results of increased familiarity and communication among the subjects included a lowering of the dimensionality seen by the subjects as necessary to encompass the stimulus space perceived as shared with the other, and perhaps the increased estimated agreement stems from this lowered dimensionality rather than from changes in the estimated ideal stimuli of the others alone.

Offhand, it is difficult to know how to provide appropriate tests of signif-

icance for this situation. Obviously, there is need to give more careful consideration to *both* kinds of information which judges may provide about objects, Task A preference and Task B similarity judgments. However, even when the psychological dimensionality of the stimulus space is known, the calculation of either exact or approximate distributions for these measures is an onerous task, as Ehrenberg has shown (4). Moreover, while these comments are directed solely to the problem of testing the significance of agreement measures, they are not unrelated to the problems of the comparison of degree of agreement across sets of judges for the same objects, and the comparison of judges across sets of objects. In either case, the same general consideration should hold: agreement in terms of Task A preference judgments reflects not only the extent to which the ideal stimulus objects of two judges are psychologically near, but also the dimensionality of the psychological stimulus spaces within which the objects are viewed by the judges in question.

SUMMARY

Within the data model introduced by Coombs, a distinction may be drawn between Task A and Task B judgments. Judgments of preference for or endorsement of items qualify as Task A, according to this model. In rank order terms, the dimensionality of the stimulus space is reflected in the number and the kind of rank orders which it is possible for a judge to produce; only when n objects require $n-1$ dimensions for description should it be possible for a judge to give all of the $n!$ permutations in order. For smaller dimensionalities, some of the permutations in order may not occur as rankings. However, the sampling theory of all available tests of agreement or concordance among judges' rankings is based upon the assumption that all $n!$ possible rankings are equally likely to occur. On the other hand, it is not particularly reasonable to assume that all stimuli which are judged under Task A instructions should require maximal psychological dimensionality (i.e., $n-1$ dimensions). The practical implication of the conflict between these two positions is that statistical tests of rank order agreement between pairs or among several judges may be seriously biased. Several simple instances of this are given as illustrations. The conclusion is that more study is needed of the psychological dimensionality of stimuli before statements about "psychological distance" between judges may be made in statistical terms.

Manuscript received: August 12, 1959

Revised Manuscript received: February 29, 1960

William L. Hays

Department of Psychology

University of Michigan

Ann Arbor, Michigan

REFERENCES

1. Bennett, J. F., "A Method for Determining the Dimensionality of a Set of Rank Orders." Unpublished Ph.D. dissertation, University of Michigan, 1951.
2. Coombs, C. H., "A Theory of Psychological Scaling." (Engineering Research Bulletin No. 34) Ann Arbor, Michigan: University of Michigan Press, 1952.
3. Coombs, C. H., "Theory and Methods of Social Measurement" in L. Festinger, and D. Katz (eds.), *Research Methods in Social Psychology*. New York: Dryden Press, 1953.
4. Ehrenberg, A. S. C., "On Sampling from a Population of Rankers." *Biometrika*, 39, 1952. pp. 82-87.
5. Hays, W. L., "An Extension of the Unfolding Technique to r Dimensions." Unpublished Ph.D. dissertation, University of Michigan, 1954.
6. Kendall, M. G., *Rank Correlation Methods*. New York: Hafner Publishing Company. 2nd edition, 1955.
7. Newcomb, T. M., "The Cognition of Persons as Cognizers" in R. Tiagiuri and L. Petrullo (ed.), *Person Perception and Interpersonal Behavior*. Stanford: Stanford University Press, 1958.

Some Factors Affecting the Behaviors of Members of Problem-Solving Groups¹

L. RICHARD HOFFMAN² AND CLAGETT G. SMITH, *The University of Michigan*

Among the various approaches to the understanding of the problem-solving process in small groups, probably the most widely accepted theory, especially among more sociologically oriented researchers, emphasizes the development of role structures in groups over time. Bales (1) has posited (and presented supporting empirical data) that, in the course of adapting to a series of problems, a group develops a set of stabilized interactions. The various functions required for accomplishing the task and for maintaining group solidarity tend to be taken on by different group members in such a way that one member will perform the leadership functions, another will initiate ideas, and still another will be well liked by virtue of his supporting other members or responding to their expressed needs. Moreover, Slater (8) has concluded that "any tendency toward role specialization . . . would increase over time, as the group became more highly 'organized' or 'structured.'"

A second approach, espoused by many "individual" psychologists, considers problem solving in groups as a pooling of individual efforts (e.g., 3). The model tends to be a probabilistic one in which the process consists of individual problem-solving activities, and in which the product is a function of the sum of the individual abilities of the group members to solve the problem. In the extreme, there is no recognition of the potential for integrative or disintegrative interaction in a group. Data supporting this view have also been reported (3).

Since both of these theories have empirical bases, it would seem that a complete theory of the problem-solving process in groups must in some way account for the success of both of these seemingly incompatible explanations, the one positing role behaviors which interact to form a group product, the other merely summing the individual efforts. The approach of this paper

¹ The research reported herein was supported by Project Michigan, a project of the University of Michigan in the field of combat surveillance, sponsored by the Department of the Army. The contract (DA-36-039 SC 52654) is administered by the Signal Corps. We wish to thank Norman R. F. Maier, the principal investigator of this sub-project, for his stimulating guidance in the conduct of the study and in the interpretation of the results. The assistance of William Hays in setting up the analysis design and the critical comments of Richard Mann on this paper are gratefully acknowledged. Parts of the analysis and write-up of the results have been supported by Public Health Service Project M-2704 from the National Institute of Mental Health.

² Now with the Department of Psychology.

posits that the particular problem requirements serve to provide the conditions in which the characteristic responses of the individual group members interact to produce a unique group adaptation to the problem. This unique adaptation will under certain conditions take the form of a role structure.

The research to be described explored the interrelationships among problem differences, characteristics of group membership, and the potential development of role structures as they affected certain behaviors of group members.

STUDY DESIGN

The data for this study have been abstracted from a study of the relative problem-solving effectiveness of groups varying in the homogeneity or heterogeneity of the personalities of the group members. On the basis of their scores on a personality inventory, students were placed in homogeneous or heterogeneous groups at the beginning of each of two semesters. Four test problems were administered to the groups on different days, and following each problem the group members rated each of the other members on six problem-solving behaviors. Using the latter measures as the dependent variable, a complex analysis of variance design was employed to determine the relative contribution of the following sources of variance: group types (homogeneous versus heterogeneous), problems, problem by group type interaction; and, within types: between groups, problem by group interaction, and between individuals within groups. We were thus able to study the effects on rated individual behavior of group characteristics (group types, between groups), of problem requirements (problems), of role structure or individual behavior consistencies (between individuals), and of certain interactions among these. The rationale of the analysis design will be discussed when the results are presented.

Subjects

Subjects for this study were the approximately 115 students in each of two semesters of an undergraduate psychology course in human relations. The study was conducted in the weekly two-hour laboratory sessions of the course. Only data from the 96 members of four-man groups in which all members were present for all four problems were used for the present analysis.

Measures

The Guilford-Zimmerman Temperament Survey was used as the personality measure (4). Ten relatively independent measures of personality are derived from this self-report, paper-and-pencil inventory: general activity, restraint, ascendance, sociability, emotional stability, objectivity, friendliness (passiv-

ity), thoughtfulness (thinking versus doing), personal relations, and masculinity. Using Kendall's tau as the measure of relationship, profile correlations were computed between every pair of students in each laboratory section (this amounted to approximately 840 correlations—15 students, taken two at a time, times eight laboratory sections). On the basis of the correlations two types of four-man groups were constructed: homogeneous groups, consisting of people with significantly positive taus, and heterogeneous groups, consisting of people with near-zero or combinations of high negative and positive taus. T-test of the difference between group types in the mean algebraic sum of taus in a group was significant at the .05 level of confidence, indicating the qualitative difference between the two types of groups. These groups were maintained for the entire semester, working together in role-playing cases and discussions. Ten homogeneous and 14 heterogeneous groups have been used in the present analysis. The homogeneous groups represented a wide variety of personality characteristics, such that in only one or two cases could the members of one group have been substituted for members of any other group.

The dependent measures of behavior for the study were obtained from post-meeting ratings of the behavior of each subject by the other three members of his group on six activities. The scales were adapted from Schutz's measures of role activities of guidance or supervision of the problem discussion, promotion of personal relations, participation in the discussion, influence over the problem solution, contribution of ideas, and attempts to be liked by the other group members.³ The exact scales used appear as Figure 1. The group members rated each person by placing the number corresponding to the person's name in one of the five boxes for each of the six scales.

It is important to note that in this method all persons are rated on all scales and ties are permitted. This method is thus different from the common technique of asking, for example, "Who participated most in the discussion?" and asks for the degree of participation of all group members.

Values from one to five were assigned to each category, as indicated in Figure 1, where five represents a great deal of the specified kind of behavior, one represents little or (in the case of guidance and promoting relations) obstructive behavior, and intermediate numbers represent intermediate levels of behavior. A rating for each subject on each scale was computed by taking the mean of his ratings by the three other members of his group.

The stability of these average measures was tested by comparing, for half of the 24 groups, the extent to which the rated differences among people in a group were greater than the discrepancies among the ratings received by

³ We wish to thank William Schutz for sending us a pre-publication copy of the questionnaire he used in his research (7).

FIGURE 1

Problem Solving Activities

Answer the following questions in terms of the way the people in your group behaved:

Guides or Supervises Problem Discussion

5	4	3	2	1
Helped very much to keep group on discussion of problem		Neither helped nor diverted group		Diverted group very much from problem discussion

Personal Relations

5	4	3	2	1
Helped very much to promote good personal feelings among group members		Neither helped nor disrupted		Caused poor personal feelings among group members

Participation

1	2	3	4	5
Hardly participated in discussion at all		Participated in discussion a moderate amount		Participated in discussion very frequently

Problem Solving

1	2	3	4	5
Had very little influence over group's problem solution		Had moderate amount of influence		Had very great influence over group's problem solution

Ideas

5	4	3	2	1
Had very good ideas and insight about problem		Had moderately good ideas and insight		Had no ideas or very poor ideas about problem

Liking by Group Members

1	2	3	4	5
Didn't care whether other group members liked him or not		Tried somewhat to be liked by other group members		Tried very hard to be liked by other group members

each subject from the three others. This was done by an analysis of variance comparing the pooled between- and within-individual data for each of the six scales. The results combined for all problems are shown in Table 1. Four of the six comparisons are statistically significant at less than .01 level of confidence. The two ratings of attempts at handling the interpersonal relations in the group, and of attempts to be liked and promoting good relations showed the least stability. On the whole, then, these ratings of the activities of group members appear to have sufficiently stable discriminability to serve usefully as the dependent measures in this study. This is true, at least, for the leadership and task activities.

TABLE 1
Reliability of Ratings of Problem-Solving Behaviors on Four Problems

Source	df	MS for					
		Guidance	Relations	Participation	Influence	Ideas	Liking
Between individuals	144	.90	.58	1.81	1.89	1.15	.61
Within individuals	384	.55	.66	.75	.78	.65	.78
Total	528						
F		1.64*	†	2.41*	2.42*	1.77*	†

* $p < .01$.

† "Within" greater than "between."

The reader should be reminded that, in spite of the discriminability of the measures, they are measures of the post-meeting perceptions of the participants in the discussions. Since during the discussion the attention of the group members was focused on solving the problem rather than observing each other's actions, these ratings could conceivably be completely unrelated to the actual activity of the group members. A more likely limitation of these measures is that the ratings were probably made relative to the general activity level of the group. In a high activity group each member could conceivably have received a lower rating on participation, for example, than would a moderately active member in a generally low activity group. These limitations of the ratings as representations of behavior should be kept in mind, but the sizeable consensus among the group members and certain other relationships between these measures and the quality and acceptance of group problem solutions (to be reported separately) suggest that these ratings would show fairly high agreement with the ratings of outside observers watching the meetings.

Problems

Since one of the major variables investigated in this research was the effects of different group problems, a short description of each problem and of the character of the problems will be presented here.

*Mined Road Problem (MR)*⁴—administered in the first session the groups met together. Adapted by Lorge to a written form from the realistic setting in which it was used for selection by the Office of Strategic Services, this problem requires the group to develop a method for crossing a road in the woods which has been saturated with supersensitive mines. Materials have been left in the vicinity which may be used in crossing the road. The instructions emphasize speed and concealment of clues from the enemy as to the group's whereabouts. The problem seems to require a certain amount of invention to develop a method for crossing and for removing the clues, and attention to a number of different aspects of the problem. There is very little in the problem to create inter-personal friction among the group members, and there is no appointed leader.

Student Assistance Fund Problem (SAF)—administered in the fourth session. This problem, developed for this research, is a four-person role-playing case with two principal parts. Each of the four participants is assigned the role of a particular type of student who is eligible, for one reason or another, to receive funds from the university which they all attend. Each one is supplied with a different reason for his eligibility, and the task of allocating funds among themselves was designed to create conflict in the group. Besides allocating the money, the group is supposed to suggest a method of determining the people who should receive money in the future. Both aspects of this problem appear to have potential for conflict based on differences in the members' situations and values. There was no appointed leader for this problem either.

Change of Work Procedures Problem (CWP in 6, p. 316 ff.)—administered in the eighth session. The conflict in this role-playing problem stems from the attempt by management, in the form of a supervisor, to change the method of work of three of his employees from rotating hourly among their jobs to each working on one job only. The supervisor introduces the subject of possible change and the group's task is to decide on a work method to be employed in the future. Here conflict to be resolved may arise in the group from the situational factors described or from attitudinal differences among the members. For this problem one person in each group was selected at random and given the supervisor's role to act as discussion leader.

⁴Irving Lorge was extremely cooperative in supplying copies of the Mined Road Problem and the scoring key.

Point Distribution Problem (PD)—administered in the tenth session. In this problem each group was given a certain number of points to be divided among the group members. The number was determined by the formula: $\text{points} = 5N - 1$ (in which N is the number of people in the group).⁵ These points were to be added to the number of points each student accumulated on his four class examinations to determine his final grade. The points could only serve to raise the grade, never to lower it. We thought that the high degree of involvement would produce a high level of interpersonal tension, but in most cases the groups awarded points as a negative function of the students' course performance, i.e., on the basis of need. Leaders were also appointed for this problem, but usually failed to perform the leadership function because they had so much at stake in the group's decision. The principal demand this problem seems to place on a group is to develop a method for resolving equitably individual needs and desires.

RESULTS

Analysis of Variance Design

Since the design of the analysis of variance in this study is so complex, a brief technical description of the design and an explanation of the components of variance examined will be given. The analysis design can be most easily understood as a two-variable (problems and groups) factorial design with "replications," where the first condition was done with homogeneous groups and the "replication" with heterogeneous groups. For illustrative purposes the analysis of variance of the guidance behaviors is presented in Table 2.⁶ A summary of the levels of significance for all six behavior analyses is presented in Table 3.

Under the column, Source, in Table 2 are listed the several sources of variance examined in this study; that is, the variables whose contributions to the variations in guidance behavior were determined in this analysis. The values in the F column corresponding to each source of variance are the results of F -tests to evaluate the statistical significance of each variable in accounting for variations in guidance behavior. The asterisk following certain F -values indicates the confidence level at which those values are statistically significant.

The significant F -value of 9.87 for "problems" indicates that there was significant variation in the average amount of guidance behavior exhibited

⁵ There were groups with other than four people working on this problem, so we needed a formula appropriate for everybody. Only four-man groups were used in this analysis to avoid bias due to size and other non-random factors.

⁶ Comparable tables for the other five behaviors are obtainable on request.

by all group members from one problem to another. In other words, guidance behavior was elicited on certain problems to a much greater degree than on others. On the other hand, the lack of significance of the F-value for "group types" reveals that the average amount of guidance behavior exhibited by members of homogeneous groups was not significantly different from that

TABLE 2
Analysis of Behavior Variance—Guidance

Source	df	MS	F
Problems	3§	277.3	9.87**
Group types	1§	59.5	2.12
Problem by type	3§	65.6	2.33
Error between groups	354	28.1	—
Within types			
Homogeneous	156	—	—
Between groups	9§	268.9	9.67**
Error between groups	147	27.8	—
Problem by group	27§	59.9	2.91**
Within groups	120	20.6	—
Between individuals	30§	35.1	2.22**
Within individuals	90§	15.8	—
Heterogeneous	220	—	—
Between groups	13§	109.2	3.87**
Error between groups	207	28.2	—
Problem by group	39§	26.7	+
Within groups	168	28.6	—
Between individuals	42§	38.8	1.54*
Within individuals	126§	25.2	—
Total	383§	—	—

+ Error variance greater.

* $p < .05$.

** $p < .01$.

§ Indicates sources of variance contributing to total degrees of freedom.

exhibited by members of heterogeneous groups. Moreover, the insignificant "problem by type" F-value shows that there were no particular problems on which the homogeneous group members showed a greater or lesser amount of guidance behavior than did heterogeneous group members. All three of these sources of variance were tested for significance using a pooled between-groups error term computed separately for each group type. The mean-square (MS) value for this error term (28.1) is listed in Table 2 under the MS column.

Since there was no possibility of matching homogeneous groups with heterogeneous groups, the remaining variations in guidance behavior must be accounted for by differences among the groups and individuals within

each group type. These analyses are shown in Table 2 separately for homogeneous groups and for heterogeneous groups. The fact that the F-values are significant for "between groups" in both the homogeneous and the heterogeneous classifications indicates that there was significant variation among the homogeneous groups and among the heterogeneous groups in the average amount of guidance behaviors exhibited; that certain homogeneous groups showed high levels of average guidance behavior while others showed lower levels, and similarly for the heterogeneous groups. The error terms for these tests were computed by pooling the "within groups" and the "problem by group" terms and the resulting significant F-values indicate a greater variation among groups of each type than among the individuals within groups.

The significant "problem by group" F-value in the homogeneous groups means that whether the average guidance behavior on any problem was high or low for a particular group, the members of each homogeneous group acted more like their fellow members than like the members of other groups. The fact that the comparable F-value for the heterogeneous groups was not significant means that the group and problem characteristics had independent effects on the guidance behavior of the members of these groups.

Finally, the significant "between individuals" F-values indicate that even within the restricted range of guidance behaviors characteristic of each group, there was a consistent ordering of individuals within groups in the amount of guidance behavior they exhibited for all four problems. Using the "within individuals" variance as the error term, the significance of these F-values indicates that there was greater variation among individuals over all the problems than there was for each individual from one problem to the next.

Problem Differences

Several consistencies are found in Table 3, of which the most important appears to be the usually underemphasized factor of problem differences. For all six behaviors the differences among problems were a significant source of variance. In meeting the varying requirements of these four problems, certain types and levels of behavior emerged regardless of other factors which were operating.

The mean ratings for each problem on the six behaviors are presented in Table 4 for all groups combined. The principal trend in this table is the relatively low level of activity in all areas on the Change of Work Procedures Problem. Since most groups reacted to the Change of Work Procedures Problem as a decision between two given alternatives, as distinct from the other three problems in which they had to invent a solution, the level of involvement may have been lower and accounted for the reduced level of

TABLE 3
Analysis of Behavior Variance
Summary Table of Tests of Significance

Source	Behaviors					
	Guidance	Relations	Participation	Influence	Ideas	Liking
Problems	**	**	**	**	*	**
Group	+	*	+	+	+	+
Problem by type	+	+	+	+	+	+
Within types						
Homogeneous						
Between groups	**	**	**	*	**	**
Problem by group	**	**	+	+	**	**
Within groups						
Between individuals	**	**	*	*	**	*
Within individuals						
Heterogeneous						
Between groups	**	**	**	**	**	**
Problem by group	+	**	+	+	+	**
Within groups						
Between individuals	*	**	+	+	+	*
Within individuals						

+ F-test *not* significant at .05 level of confidence.

* F-test significant at .05 level of confidence.

** F-test significant at .01 level of confidence.

activity. The relatively high level of task-directed behaviors (guidance, participation, influence, ideas) on the Mined Road Problem and of social-emotional behaviors (personal relations and liking) on the Point Distribution Problem were expected from the emphasis on quality and acceptance, respectively, for these two problems. The Student Assistance Fund Problem elicited generally high rates of activity on the influence and ideas scales, but very low levels of social-emotional behavior. The groups tended to concentrate their attention on the part of this problem in which they were asked to recommend a method for the future allocation of funds rather than on the distribution of the money among themselves, which was easily reconciled in terms of need. Thus each problem showed a relatively distinct pattern of behaviors, emphasizing the different requirements each of the four problems placed on the groups.

Group Types

The direct effects of the homogeneity or heterogeneity of the group membership on these activities appear to have been minimal (Table 3). The only statistically significant variance attributable to this group characteristic was

the greater promoting of personal relations by the members of homogeneous groups (means were 3.85 for the homogeneous and 3.71 for the heterogeneous groups). There were no significant differences between the group types on the other five behaviors. Nor were there any direct interactional effects of group types and problems. Whatever behaviors were evoked by different problems emerged, on the average, to a similar degree in both types of groups.

Within each group type, moreover, relatively large and significant differences were found among groups on all six behaviors (Table 3). Each group developed a distinct and characteristic level of activity on each behavior in

TABLE 4
Problem-Solving Behaviors on Different Problems

Problems *	Behaviors					
	Guidance	Relations	Participation	Influence	Ideas	Liking
Mined road	4.22*	3.83	3.97	3.59	3.66	3.42
Student assistance fund	4.04	3.60	3.74	3.59	3.72	3.17
Change of work procedures	3.83	3.65	3.51	3.33	3.47	3.31
Point distribution	4.16	4.00	3.56	3.46	3.69	3.50

* Problems are listed in the order in which they were administered during each semester.

* Mean ratings for 96 group members in 24 groups. Ratings on any behavior may be compared for different problems, but ratings for different behaviors are not comparable.

addition to the way all groups responded to each problem. Whether these differences among groups arose from initial differences in structural characteristics other than homogeneity-heterogeneity of personality, from the groups' differential adaptive processes, or from the interaction of these remains to be determined.

Differential effects of homogeneity-heterogeneity did appear, however, in certain analyses within group types. Significant problem-by-group interaction variance was found in the homogeneous groups on four behaviors (but not on participation or influence); but only promoting personal relations and attempting to be liked were significant in the heterogeneous groups. Examination of the group means indicates that rather than intensifying the level of behaviors found to be characteristic of each problem, each homogeneous group reacted in a distinct fashion to each problem, sometimes higher than the problem mean and sometimes lower. The same groups were seldom consistently higher or lower. It would appear that the requirements of each problem had different effects and placed different strains on each group. The findings also indicate that whatever the effect of a problem on a group, the members of each homogeneous group reacted more like their fellow group members than did the members of heterogeneous groups, especially for the task-

directed behaviors. In the social-emotional area the actions of the members of both homogeneous and heterogeneous groups were similar to those of the other members of their groups, suggesting that different group levels of emotional exchange were developed for each problem.

Role Structure and Individual Differences

A similarly consistent finding which modifies the last interpretation appears in the comparison of the between-individuals-within-groups-within-types variance (Table 3). This measures the degree to which there was an ordering among the individuals in each group across problems in the amount of each behavior expressed. In the homogeneous groups the between-individuals variance was significant for all six behaviors, whereas it was significant for only three behaviors in the heterogeneous groups. Thus, although the members of homogeneous groups, on the average, were more likely to have behaved in a distinct and characteristic manner than were the members of heterogeneous groups, the small differences in behaviors among the group members tended to be more stable in the former than in the latter groups.

The significance of the between-individuals-within group variance in the homogeneous groups on all six behaviors suggested that, in accordance with the more sociological point of view, a role structure might have developed in these groups. Certain group members might have more often behaved in one way while others behaved in a different way.

A measure was computed of the degree of behavioral differentiation among the members of each group. The absolute deviation of each member's rating from his group's mean rating was computed on each behavior and summed over all behaviors for each problem. The higher this measure the more differentiated the members were seen to have been with respect to the behaviors rated and, therefore, the higher the degree of role-structure achieved. If the role-structural hypothesis is correct, this index would be expected to become greater the more times the groups interacted.

The mean differentiation indexes are presented for each of the four problems, separately for the homogeneous and heterogeneous groups, in Table 5. The problems are listed in the order they were administered. It will be seen that, contrary to the hypothesis, the size of this index decreases from the early-administered problems to the later ones. The drop is statistically significant for the homogeneous but not for the heterogeneous groups, although it drops somewhat in the latter groups also.

Relations Between Behaviors

Apart from the results of the analyses of variance, the intercorrelations among the six behavior ratings are of interest in themselves. The intercorrela-

TABLE 5
Indexes of Role Differentiation

Group type	No. of Groups		Problems				F
			MR	SAF	CWP	PD	
Homogeneous	10	Mean	.45*	.37	.35	.27	3.58*
		Stand. Dev.	.158	.116	.108	.106	—
Heterogeneous	14	Mean	.46	.46	.39	.36	2.39
		Stand. Dev.	.145	.155	.123	.102	—

* The higher the number the greater the differentiation.

* $p < .05$.

tions among the mean ratings on the four problems for all groups combined are presented in Table 6. High correlations appear between all the scales except attempts to be liked (in sociometric data, which will be reported separately, people who tried to be liked were, in fact, rated as being liked by the other members of their groups, so that this measure may be considered as somewhat comparable to liking measures used by others). Thus the separation of task and social-emotional functions was found again here as in other studies. Guidance appears to be a general activity rating that embodies both or either of these two functions, as seen in its high correlations with the other five behaviors and in the fact that it correlates highest of the six with mean activity level. Ideas and influence form a somewhat separate constellation of task-directed behaviors. And the high correlation between relations and liking suggests a set of social-emotional behaviors.

Comparable correlational matrices, computed separately for the homogeneous and heterogeneous groups, showed remarkable similarity. The one major difference between the correlations was the significantly higher correlation between participation and influence in the heterogeneous groups ($r = .84$) than in the homogeneous groups ($r = .45$). Although both correlations are significantly different from zero, participation in the discussions

TABLE 6
Intercorrelations of Problem-Solving Behaviors

	Guidance	Relations	Participation	Influence	Ideas	Liking	Mean Level of Behavior
Guidance	—	.60*	.57	.68	.59	.42	.88
Relations		—	.31	.45	.56	.57	.77
Participation			—	.68	.53	-.05	.68
Influence				—	.70	.21	.83
Ideas					—	.29	.81
Liking						—	.53

* For $N = 96$: $r = .20$ is significantly different from zero at the .05 level of confidence.

in homogeneous groups did not ensure as much influence over the final decision as it did in the heterogeneous groups. Much of the behavior rated as participation in homogeneous groups probably did not contribute radically new ideas which altered the basic solutions to the problems, whereas the contributions of members of heterogeneous groups were more effective in this respect. Except for this difference, the separation of functions described above for the combined groups was maintained for each of the group types separately.

DISCUSSION

The most interesting aspect of the results reported here is the importance of the problem requirements in determining the behaviors of the group members. Each problem used in this study appears to have elicited certain types of reactions from the group members, and these reactions differed from problem to problem. The problem requirements appear to have provided a stimulus to which each group member responded. This conclusion, if it is supported by direct observational data, indicates the importance in all research on small-group process of identifying the nature of the task given to the group,⁷ or at least of describing the problem in sufficient detail that the reader may draw his own conclusions about its nature. The common practice of not identifying the problem used or of stating merely that "a problem of the human relations variety was used" deprives the researcher on group process of a great deal of important and necessary information for his understanding of the results of the experiment reported.

Two possible dichotomies for classifying problems are suggested by the data. The first is the distinction made by Maier between problems of *quality* and those of *acceptance* (5). Solutions to quality problems may be evaluated by some criterion external to the group; solutions to acceptance problems cannot be so evaluated and rest on their ability to satisfy the group members. The Mined Road Problem was the purest quality problem of the four and showed the highest rates of activity on the guidance and participation scales, keeping the discussion problem-centered and the members active. The Point Distribution Problem was the purest acceptance problem of the four and the members were most active in improving personal relations and attempting to be liked.

A second dichotomy which seems useful is that between *decision-making* and *creative* problems. The majority of groups responded to the Change of Work Procedure Problem as a choice between two given alternatives and

⁷ Unfortunately little systematic attention has been given to the classification of group tasks according to the requirements they demand from a group, so that this type of description is now impossible to do.

usually chose one or the other. For the other three problems the groups had to create a solution from the information, since none was given in the problem. It is thus significant that the Change of Work Procedures Problem showed the lowest rate of activity on all four task-related behaviors—guidance, participation, influence and ideas. It would seem that creating a solution may require a greater expenditure of energy devoted to solving the problem than does merely choosing among given alternatives.

The present findings also provide perspective on the place of the "individualistic" and "role theory" conceptions of group process. Although characteristic individual differences appeared on certain behaviors in both homogeneous and heterogeneous groups, as evidenced by the significant between-individual variances found, the proportion of variance accounted for by these differences was small as compared to the contributions of problem and group differences. To understand a person's behavior in a group the problem requirements and the group characteristics must first be considered and then the personal attributes of the individual may be looked at within this over-all framework.

A similar comment may be made about the role structural explanation since the data testing it were the same. But this explanation is even further qualified by the generally negative finding that the indexes of role differentiation showed a disintegration of role structure over time, rather than a development. This contrasts with Slater's finding that in his groups the role structures become better defined as the groups worked together (8). It is conceivable that these differences may be due to any or all of three factors. The first is that his groups met only four times, in contrast to the ten meetings held by our groups. The indexes of role structure for the Student Assistance Fund problem, which was administered in the fourth group meeting, show little change from the indexes on the Mined Road Problem, the first problem of the semester. By the tenth meeting, however, when the Point Distribution Problem was given, the differentiation indexes had dropped sharply in both types of groups. Thus, although groups probably show stable role structure in the first four meetings, the members may come to feel less constrained in their behaviors as they continue their interactions.

Another possible cause of the differing results from the two studies is the type of problems used by Slater. Unfortunately the only description of the problems in the article is: "... an administrative problem, which the subjects were asked to solve as a group, assuming the role of an administrative staff to the central authority in the case under discussion. They were given 40 minutes to discuss the case and decide (a) why the persons involved in the case were behaving as they did, and (b) what the central authority should do about it" (8). These cases probably come closest to the Mined Road Problem and to the advisory part of the Student Assistance Fund problem

used in the present study, since the subjects were essentially uninvolved in the outcome and had to create solutions. The role structure indexes were highest for these two problems and, to the extent that the emergence of role structures may be related to the type of task, his groups developed such structures. Furthermore, since the problem requirements may have remained fairly constant for the four problems used by Slater, the groups are likely to have become more efficient over time, and more frequently and quickly looked to certain people for guidance in solving this type of problem.

The third possible cause of the contrasting results rests in the difference between considering only the highest-ranked man in each group, as Slater did, and determining the degree of structure in the entire group, as was done in the present analysis. Although this is probably the least likely explanation of the three here suggested, the reader should be aware of it as a possibility. Subjects are able reliably to rank-order the other members of their groups on the activities we are concerned with here (although we have doubts about the adequacy of their self-rankings). In many groups (in our data this became more true as time passed) the difference in activity rate between the first- and second-ranked men is negligible. Most of the low status-consensus groups in Slater's study were probably of this type. Probably, however, many of the high status-consensus groups were also. From the usual discussion of role specialization in groups one obtains a picture of one member being an idea man, another member the one who is liked, a third one a mediator. This image is perpetuated by Slater's type of analysis, but the question remains as to how much each member, other than the top-ranked one, engages in each of these functions. Our data would suggest that the members may restrict their activities less and less to a single function, and, as the group gets older, more often act as they perceive the situation requires.

These three questions are certainly amenable to empirical answers. A study is needed in which groups constructed with a range of differing characteristics meet many times to solve a variety of problems, and for which both ratings and direct observations of the members' behaviors are obtained. The data on the behaviors should then be analyzed, in terms both of the top-ranked members and for all the group members, for the development or disintegration of role structures over time and for different problems.

Finally a few words should be said about the results of the homogeneity-heterogeneity of personality classification. Three major differences emerged between the two types of groups. First, on the average the members of homogeneous groups engaged in more promotion of personal relations than did those in heterogeneous groups. Although unexpected, this finding lends support to the idea that the members of homogeneous groups more than the members of heterogeneous groups expressed agreement with other people's ideas and praise for "good thinking," rather than searching for new ideas which

caused the former groups to be inferior problem solvers.⁸ The possibility should not be ruled out that in many of the homogeneous groups the greater attention to improving personal relations was a response to a high level of interpersonal tension *caused* by their personality similarities. Cohen (2) has shown that the interactions of people with certain types of similar personalities is actually unpleasant.

The second difference between the two types of groups was the somewhat greater frequency of significant problem-by-group interactions among the homogeneous groups than among the heterogeneous groups. This measures the degree to which the members of each group acted more like each other in responding to a problem than like members of other groups, and indicates that these responses were different in each group for different problems. Although groups of both types appear to have developed characteristic levels of social-emotional activity to each problem, the members of homogeneous groups also produced significant problem-by-group interactions on the guidance and ideas scales. Thus, the members of homogeneous groups showed somewhat more similar levels of activity in their responses to the problems than did the members of heterogeneous groups.

In contrast to this last finding, which was expected from the basic similarity of personality of the members of the homogeneous groups, the third difference between the two types of groups showed up in the significant between-individuals differences in the homogeneous groups on all six behaviors. It appears that, within the narrower range of activity levels within each homogeneous group, the members more often maintained a consistent ordering than did the members of the heterogeneous groups in a usually somewhat wider range. These differences are difficult to understand. They could have occurred because there was a greater tolerance for divergent behaviors, given a basic similarity of predispositions; because individual characteristics other than those used to assign people to groups were operating; or because through historical chance each person initially behaved somewhat differently and few forces developed to alter this pattern. In interpreting the results for the homogeneous groups, recall that they are a function of similarity of personality among the members of each group and not a function of any particular personality characteristics.

In both types of groups large and unexplainable between-group differences were found. These suggest the emergence of a group norm of activity and characteristic level of responding by the group to every problem. These large between-group differences, however, were probably most subject to any response biases which might have existed within any group, and most clearly need verification by direct observational procedures. It would be interesting

⁸ Data showing the relatively superior problem-solving effectiveness of the heterogeneous over the homogeneous groups will be reported separately.

to determine also what factors caused the emergence of these within-group consistencies and whether the same factors operated in homogeneous as in heterogeneous groups.

These findings form the basis of an approach to the understanding of small-group process, the basic premises of which are: (a) different externally imposed problems place different demands on a group; (b) group members respond individually to problem demands according to their unique characteristics; (c) differing individual reactions, either individually or in interaction with the reactions of the other group members, may result in tendencies to perform specialized role functions, but the tendency of members to perform different functions is not a matter of playing or not playing a role, but rather a matter of performing more or less of the function; (d) reactions in interaction produce a variety and level of group behavior which is more than and different from a mere summing of individual potentials, but which is characteristic of the group as a whole in each problem situation; (e) the more similar the attributes of the individuals composing a group the more likely they are to be similar to each other in their responses to problems in a group situation. This approach provides a framework for more intensive research on the group problem-solving process directed toward the testing of the present findings with direct observational data and the discovery of the problem, group, and individual characteristics which interact to produce particular forms of member behavior.

SUMMARY

Four-person groups, composed according to the homogeneity or heterogeneity of their members' personalities, provided data for an examination of some of the determinants of the behaviors of members of problem-solving groups. The groups, established in the laboratory sections of an undergraduate course, worked on four different problems at different times during the semester. After each problem-solving session the group members rated each other on six behaviors: guidance or supervision of the discussion, promotion of personal relations, participation in the discussion, influence over the decision, contribution of good ideas, and attempts at being liked by others. These ratings served as dependent measures for a factorial analysis-of-variance design with sources of variance for: problems, group types, between groups, between individuals, and for a number of interactions among these.

Differences among the problems were a major source of variance on all six behaviors. Each problem showed its own characteristic configuration of behaviors which seem to have been elicited by the requirements of the problem. The usefulness of classifying problems in terms of the dichotomies *quality* versus *acceptance* and *creative* versus *decision-making* is discussed in terms of the different kinds of behaviors and levels of activity evoked.

The data are also discussed in terms of the questions they raise for the adequacy of "individualistic" and "role theoretical" theories of group process, and in terms of the suggestions they have for further research on these problems.

Homogeneous groups engaged more in promoting personal relations than did heterogeneous groups, but the differences between the group types were not statistically significant for any other behavior. Differences were found within group types, however. Problem-by-group interaction and between-individual variances were more often significant in homogeneous than in heterogeneous groups.

Large between-groups differences were also found in both types of groups on all six behaviors, but the data were insufficient to determine what factors caused these differences. Another question left unanswered by the data was whether the same factors caused the differences among the homogeneous as among the heterogeneous groups.

Basic premises for an approach to an explanation of the group problem-solving process are presented, based on these findings. These are stated generally as a guide to research using direct observational methods to identify problem, group, and individual characteristics and to study their interrelationships using a variety of problems over an extended period of group life.

Manuscript received: July 13, 1959

Revised Manuscript received: December 14, 1959

L. Richard Hoffman
Department of Psychology
University of Michigan
Ann Arbor, Michigan

REFERENCES

1. Bales, R. F., "The Equilibrium Problem in Small Groups," in T. Parsons, R. F. Bales, and E. A. Shils, *Working Papers in the Theory of Action*, Glencoe, Illinois: The Free Press, 1953.
2. Cohen, A. R., "Experimental Effects of Ego-Defense Preference on Interpersonal Relations," *Journal of Abnormal and Social Psychology*, 1956, 52, 19-27.
3. Ekman, G., "The Four Effects of Cooperation," *Journal of Social Psychology*, 1955, 41, 149-162.
4. Guilford, J. P., and W. S. Zimmerman, *The Guilford-Zimmerman Temperament Survey*, Beverly Hills: Sheridan Supply Co., 1949.
5. Maier, N. R. F., *Principles of Human Relations*, New York: Wiley, 1952.
6. Maier, N. R. F., *Psychology in Industry* (2nd Edition), New York: Houghton Mifflin, 1955.
7. Schutz, W. C., *FIRO: A Three-Dimensional Theory of Inter-personal Behavior*, New York: Rinehart, 1958.
8. Slater, P. E., "Role Differentiation in Small Groups," *American Sociological Review*, 1955, 20, 300-310.

A Ritual Pollution Scale for Ranking Hindu Castes¹

PAULINE MOLLER MAHAR, *University of Arizona*

In his recent review of the past studies of caste ranking in South Asia,² McKim Marriott concluded:

Studies of social life in South Asia have reached an impasse on the problem of measuring relative caste ranking. . . . The problem of measuring caste rank is crucial for many kinds of social research in India, Pakistan, and Ceylon (9, p. 1).

Marriott himself then offered a technique he has used with considerable success for the study of ranking and of the varied criteria determining caste rank. The present report was prompted by the problem Marriott pointed up. It offers a method for measuring caste ranking which is more limited in scope, but, hopefully, more refined than those usually used in the study of caste hierarchy. Results based on its application suggest a rather high degree of consensus in the village studied concerning caste ranking and ritual bases of rank.

The method offers the investigator: (a) considerable control over the criteria for ranking, (b) a means for systematic analysis of responses from many informants according to an objective test for consensus, (c) a means for pinpointing dissensus as well as consensus, and (d) an objective means of replication of the study in the same locality.

Through use of this method, it was possible to reveal a hierarchy in which 18 out of 22 castes could be fairly precisely located in terms of 15 different

¹ An earlier version of this paper was read at the fifty-seventh annual meeting of the American Anthropological Association, November, 1958, as part of the symposium, Caste in India, and was published with the symposium in *Man in India*, 1959, 39, 127-147. The current version presents some new material and substantial rewriting for the audience of *Sociometry*.

The ritual distance study was undertaken during the winter of 1955-56 in Khālāpur, a village in northwestern Uttar Pradesh, where I carried on various field projects for 16 months between October, 1954 and May, 1956 as a post-doctoral fellow of the Cornell University India Program. The paper was first written in 1958 when I was a grantee of the American Philosophical Society. Such support is gratefully acknowledged.

For help on the ritual distance study I owe special thanks to Miss Usha Bhagat and Miss Saubhagya Taneja who were interviewers, and to Dr. John Gumperz, a linguist, who helped Miss Bhagat and me frame the interview schedule. I should also thank Mr. J. Michael Mahar and Dr. McKim Marriott who encouraged me to analyse these particular data, Edward B. Harper, David Mandelbaum, Morris E. Opler, Martin Orans, James Silverberg, Wayne Thompson and the staff of *Sociometry* for many helpful criticisms and suggestions.

² The studies reviewed by Marriott included 3, 4, 6, 8, 10, 14 and 17 in References.

ranks. Those castes upon whose ranking there was disagreement could be specified. The broad classes of castes as perceived by the villagers were suggested by the ranking evidence.

THE RITUAL DISTANCE INTERVIEW

In an effort to rank the Hindu and Muslim castes of Khālāpur, a village in western Uttar Pradesh (see appendix for the list of castes), I limited my concern to one class of criteria for ranking, that of ritual purity and pollution.

The use of the word 'ritual' is akin to Durkheim's concept of the 'sacred' and to the Polynesian 'tabu.' Radcliffe-Brown explains:

A ritual prohibition is a rule of behaviour which is associated with a belief that an infraction will result in an undesirable change in the ritual status of the person who fails to keep to the rule (12, pp. 134-135).

What has been called 'the Hindu Pollution Concept' (18, p. 63) has been considered by a number of students to be the crucial characteristic of the Indian caste system. H. N. C. Stevenson says:

... one thing seems certain, that it is from ritual rather than secular status, and from group rather than personal status, that the caste system derives its unique consistency and viability (18, p. 47).

Stevenson reviews some of the proponents of this point of view. These include S. V. Ketkar, J. H. Hutton, Sarat Chandra Roy, G. S. Ghurye, and M. N. Srinivas. Srinivas, for example, has said:

This concept (of pollution) is absolutely fundamental to the caste system, and along with the concepts of *karma* and *dharma* it contributes to make caste the unique institution it is (16, p. 26).

In comparing caste in Africa and India, S. F. Nadel has concluded:

In Bharat India the conception of spiritual purity and perfection dominates the picture, so that caste differentiation in turn expresses the ideas of pollution, contagion, and segregation between people of different degrees of perfection. In Africa the dominant conception is more pragmatic and in a sense materialistic (11, p. 22).

Within the category of ritual purity and pollution, I tried to learn about the norms relevant to the interactions symbolizing inequality of ritual status in dyadic inter-caste relations. Such customs include those relating to touch, eating, drinking, and smoking when members of two different castes are in contact. If consistent rules of inequality govern all paired relationships, the combination of all pairs results in a consistent caste hierarchy for a village or locality.

The beliefs concerning the ritual purity or impurity of various castes'

occupational, dietary and marriage customs may be reflected in such a tactile-consumption ritual hierarchy. However, they are not directly tapped as criteria for ranking in the Ritual Distance Interview, the technique which I shall describe and illustrate in this paper. Other dimensions of stratification, those which H. N. C. Stevenson has called 'secular' (18, p. 63), especially economic and political power of various kinds, may also be reflected in this ritual hierarchy, but they were not *directly* involved.

The Ritual Distance Interview adapted ideas from the Bogardus Social Distance Scale (2)³ and from Blunt's ritual ranking of castes of the United Provinces (1), and combined these with the Guttman scaling technique for ranking objects according to a single consensual attitude dimension (7, 19). Recently, the Guttman scale has been elaborated into types of scales by Matilda Riley, John Riley, Jackson Toby and their colleagues (13). The types they have called the "object scale" and the "group scale" may be combined into a technique applicable to the study of ritual ranking of Indian castes.

After considerable preliminary investigation of the ritual practices of the Khālāpur villagers, questions were framed appropriately for a Guttman scale model. With the help of a linguist, John Gumperz, we tried to state questions in language natural to natives of Khālāpur. The questions of the Ritual Distance Interview appear in the Appendix.

The items were formulated in such a way that, if the attitudes were unidimensional by the test for a Guttman scale, each item would represent a different point along a single continuum of ritual purity and pollution. Acceptance of any particular item of the Guttman scale of ritual practices would consistently be associated with acceptance of all items representing behaviors permitted to persons at a greater ritual distance.

THE SCALE-PICTURE

Figure 1 gives an example of such a ritual distance scale. This is the scale-picture of the 273 responses of Manbhi, a Brahman woman. She was asked 13 questions about 21 castes. (Her own caste was omitted.) We can see from looking at the top of the picture, that she would allow members of the Rajput, Merchant and Water-carrier castes to perform ten of the 13 ritual actions. At the bottom of the picture, it can be seen that she would not allow members of the Chamar Weaver, Shoemaker, Untouchable Agricultural Laborer or Sweeper castes to do any of these things. Between these two extremes, the other castes can be regularly ranked according to the action of

³ For a summary of the studies of "social distance" in the United States, see Simpson and Yinger (15). For another study in India using the Bogardus Social Distance Scale, see Guha and Kaul (5).

FIGURE 1

Scale-Picture of Responses Indicating Ritual Distance for Manbhi, Brahman Caste, Female, 40

Item	13*	10	1*	2	7*	11	4*	8*	6	9*	12	3	5	Score
Can touch our children														
Can accept dry, uncooked food														
Can touch me														
Can sit on our cot														
Can smoke bowl of pipe														
Can take water from his hand														
Can touch our brass vessels														
Can accept fried (<i>Pakka</i>) food from him														
Can smoke our pipe														
Can accept boiled (<i>Kachcha</i>) food from his hand														
Can touch our water vessel														
Can come on our cooking area														
Can touch our earthenware vessels														
CASTES:														
Rajput	x	x	x	x	x	x	x	x	x	x				10
Merchants	x	x	x	x	x	x	x	x	x	x				10
Water-Carriers	x	x	x	x	x	x	x	x	x	0	x			10-11
Goldsmiths	x	x	x	x	x	x	x	x	x					9
Genealogists	x	x	x	x	x	x	x	x	x					9
Barbers	x	x	x	x	x	x	x	x						8
Goosaali	x	x	x	x	x	x								6
Shepherds	x	x	x	x	x	x								6
Carpenters	x	x	x	x	x	x								6
Potters	x	x	x	x	0	x								5-6
Washermen	x	x	x	x	x									5
Grainparchers	x	x	x	x	x									5
K. B. Weavers	x	0	x	x										3-4
Joogiis	x	0	x	x										3-4
Mus. Rajputs	x	x	x											3
Oilpressers	x	x												2
Miraassiiis	x	0	x											1-2-3
Ch. Weavers														0
Shoemakers														0
Ag. laborers														0
Sweepers														0

* Starred items are those used in the final scale derived from the responses of eighteen interviewees.

No. Deviant Responses: 5

Total Responses: 273

least ritual distance permitted. Thus, members of those castes (i.e., Washermen and Grainparchers) whom Manbhi will allow to smoke the bowl of the pipe (item 7), but to whom she *prohibits* actions of *lesser* ritual distance, she *permits* actions of *greater* ritual distance. They can sit on the cot, they can touch her and her children, all items indicating greater ritual distance. Manbhi's responses scale according to the Guttman test since, for any behavior permitted, she also allows members of a caste to whom that right is granted to do everything of *greater* ritual distance.

THE RESPONDENTS

Ritual Distance interviews for 18 respondents from 11 different castes, were analyzed. (See Table 1 for their distribution.) These respondents were

TABLE 1
Caste and Sex of Eighteen Respondents

	Male	Female
Rajput	1	0
Brahman	1	1
Merchant	2	0
Water-carrier	1	2
Carpenter	1	2
Potter	0	1
Barber	0	1
Washerman	1	0
Genealogist	0	1
Chamar	1	1
Sweeper	1	0
	$\bar{9}$	$\bar{9}$

part of a pre-test for a Ritual Distance Survey undertaken in Khālāpur. Interviewees for the pre-test were selected in three different ways. Ten of them were friends or acquaintances we had made in the course of specialized studies of caste communities or of calendrical rites. Five were young men with eight or ten years of schooling whom we interviewed for the first time when hiring interviewers to carry out the survey itself. Three persons we knew only through the pre-test. They were selected because they were at home when we wanted a person of a particular caste, or, in one case, the respondent fell into a random sample taken for the survey and was interviewed reliably. The large-scale survey itself was never completed, I might mention, because we found the schoolboys, the only potential group of interviewers, to be too immature to carry out such a study reliably.

None of the interviewees can be considered to be a village leader. All of

the women and one of the men are completely illiterate. The other eight men vary in schooling from four to ten years.

The 18 interviews used in this analysis were done by Miss Usha Bhagat and Miss Saubhagya Taneja. I myself was present during eight of the interviews. Each interview lasted about one hour. Private interviews were obtained with 14 of the respondents. In the four other cases one or more family members or neighbors were present and made comments. Since the respondent was the eldest person present in three of the cases, however, there was no need for him to defer to the others, and the fourth, a Chamar woman, was outspoken in her disagreements with some of the remarks made by those around her.

THE PURITY-POLLUTION CONTINUUM

In the analysis of the ritual distance interviews of these 18 respondents, I found that the responses of all could be diagrammed in the Guttman-scale fashion.

On the basis of the number of actions permitted each caste by a respondent, as recorded in his scale-picture, it is possible to assign each caste a score. The scores derived from Manbhi's responses are listed in the right hand column of Figure 1. From these scores, we can arrive at the hierarchy of castes representing Manbhi's ritual ranking of 21 castes of Khālāpur. In the same way, it is possible to derive ritual caste rankings from the scale-pictures of the other 17 respondents. By comparison of the 18 individual hierarchies it is possible to derive the ritual ranking of castes of Khālāpur village as a whole.

As might be expected, respondents tended to vary the order of the items as revealed in their respective scale-pictures. One item would appear to signify greater ritual distance for some respondents than for others. However, all respondents ordered six of the items in an almost identical sequence, items 13, 1, 7, 4, 8, and 9. (See questionnaire in Appendix.)

Five of the other items were not placed along the continuum in precisely the same positions by all informants, although there tended to be considerable agreement about general location.

We may ask why the purity-pollution items fall in this particular sequence. On the basis of Stevenson's conceptualization (18), the principles involved include: (a) the greater seriousness of internal pollution (taking polluted food or water) over external pollution (touch), (b) the conductivity of vessels, with the further qualification that *porous* earthenware vessels may be internally pollutable while hard metal vessels are only externally pollutable, and (c) the protective nature of such products of the cow as milk, ghī (clarified butter) and dung, so that foods cooked in milk or ghī, for example, are less easily polluted than foods cooked in water.

Degrees of seriousness of pollution are suggested by these village respondents not only in terms of the numbers of castes allowed to perform the various acts, but also in the spontaneous comments made by ten of them regarding means of purification. While defilement from touch of the person may be dispelled through bathing and washing one's clothes, a polluted brass vessel may be purified only through scouring it with hot coals, and a polluted earthenware vessel must be discarded entirely. No means of purification after taking polluted food or water were offered.

DEVELOPMENT OF THE OVER-ALL RANKING

After deriving the ordinal relationship of the items of interaction, a scoring system ranging from '6' to '0' was constructed with the six stable items of

TABLE 2
Scoring from Six Items of the Purity-Pollution Continuum

Score	Items Acceded To					
0	—					
1	13					
2	13	1				
3	13	1	7			
4	13	1	7	4		
5	13	1	7	4	8	
6	13	1	7	4	8	9

the purity-pollution continuum. In such a system, each score represents an identical combination of responses to items by all informants. Thus, a score of '3' always is to be understood to be a combination of items 13, 1 and 7. The scoring in terms of the items acceded to appears below in Table 2.

To have had a *perfectly* consistent over-all ranking, all eighteen respondents should have given the same castes the same scores. This was not the case. Individuals varied in the combinations of castes given each score. The way in which it was possible to develop an over-all ranking which was consistent with, though not identical for, all eighteen hierarchies, was to focus upon the *cutting points* in the scorings, the points in a hierarchy where the score of '6' ended and '5' began, where '5' ended and '4' began, and so on. Rankings were made on the basis of differentiations which were a function of the cutting points.

To illustrate how a decision on ranking was made, let us compare the Merchants and Goldsmiths. Twelve informants considered Merchants and Goldsmiths to be equal. They gave the two castes the same Ritual Distance scores. Four individuals, however, gave the two castes different scores. All four of these gave the Merchants the higher score. We may assume further

that the two Merchant informants whose cutting points come at their own caste's rank would have given their own caste the higher score. So it is on the basis of six individuals' differentiations that Merchants are ranked higher than Goldsmiths.

By this method of comparison of all pairs of castes, there emerged the ranking in the left column of Figure 2. This figure also presents the caste scorings and rankings derived from the responses of six of the respondents, to illustrate the form of the 18 lists which were compared.

No caste was included in the over-all hierarchy upon whose rank as given here more than three respondents disagreed. The Genealogists, Shepherds, Water-carriers, Muslim Rajputs and Shoemakers are the castes for each of which two informants deviate from the over-all rankings given here, and the Mirāssī Beggars are the one caste in the ranking for which three persons deviated. Two or three disagreements seemed small in number, for even with these, it meant that any particular ranking was consistent with the rankings of more than four-fifths of the informants. The standard of four-fifths agreement is, of course, an arbitrary one.

Between each pair of castes in Figure 2, to the right of the hierarchy with the names of the castes, the numbers of respondents determining each differentiation are listed. With a larger number of respondents it might be possible to develop *intervals* of ritual distance on the basis of the numbers of respondents differentiating the various pairs of castes. Only one person among these interviewees, for example, ranks the Brahmans above the Rajputs, and only one person rates the Rajputs above the Merchants. Singular differentiations such as these would suggest that the castes are just about equal in rank as measured by this particular Ritual Distance Scale.

The larger number of differentiations in rank would suggest a widely-shared perception among the villagers of a definite gulf separating the contiguous castes so distinguished. On the basis of the numbers of differentiations made in these interviews, we may infer that the major blocs of castes perceived in Khālāpur are the clean Hindu castes, the Muslim castes, and the Untouchable castes. The number of respondents separating the lowest clean Hindu caste, the Shepherds, from the highest Muslim caste, the Muslim Rajputs, is twelve. The number differentiating the lowest Muslim caste, the Mirāssīs, from the highest Untouchable caste, the Shoemakers, is twelve. It is noteworthy that there is no such well-agreed-upon break between a bloc of "twice-born" castes at the top and the *Shūdrā* or servant castes in the middle of the hierarchy. The aristocracy of Khālāpur is not well-marked, at least by this method of ranking.

The seven castes which it was not possible to rank, given the requirement that there must be four-fifths' agreement on each ranking, are of two kinds. First, there are three castes, the Chamar Weavers, Potters and Washermen,

FIGURE 2

*Eighteen Hindu Villagers' Scoring and Ranking of Fifteen Castes for
Ritual Purity with Six Illustrative Scorings and Rankings*

Castes in order of Over-all ranking	Number of Respondents Placing Caste Above the Succeeding Caste (N=18)	Scoring and Rankings by Six of the Eighteen Respondents					
		f-40 Brahman	M-20 Rajput	F-40 Barber	M-30 Washer	F-55 Carpent.	M-17 Sweeper
BRAHMANS	1	—	6	6	6	6	6
WARRIORS (Rajput)	1	6	—	6	6	5	6
MERCHANTS (Baniyās)	6	6	6	6	6	5	6
GOLDSMITHS (Sunārs)	7	5	6	6	6	5	6
GENEALOGISTS (Bhāts)	3	5	6	5	6	④	6
BARBERS (Nāis)	3	5	4	—	6	5	6
WATER-CARRIERS (Jhīnvars)	7	5	4	4	5	5	6
GRAINPARCHERS (Bharbbūjās)	6	3	4	4	5	4	6
SHEPHERDS (Garariyās)	12	3	4	4	4	3	6
MUSLIM RAJPUTS	10	1	2	2	3	3	6
OIL PRESSERS (Telis)	6	1	* ④	2	3	2	4
BEGGARS (Mirāsī Dūm)	12	1	2	③	2	2	4
SHOEMAKERS (Jatīyās)	3	0	1	0	1	1	4
AG. LABORERS (Chamārs)	1	0	0	0	1	0	4
SWEEPERS (Chūrhās)	0	0	0	0	1	0	—

* Numbers circled deviate from the over-all ranking. Cutting points.

whose approximate positions in the hierarchy can be located quite accurately. Most respondents placed the Chamar Weavers, for example, just above, equal to, or just below the three Muslim castes.

Secondly, there are the castes which have marked religious identification:

the Kabir Panthī Weavers, the Jōgīs, the Gōsāīs, and the Carpenters (who style themselves Dhīman Brahmins). Their positions are less easily located. These are castes which are ranked high by respondents who take the religious identification more seriously than others do.

The Single Informant and the Caste Hierarchy

The caste hierarchy we have been able to develop for village Khālāpur is based upon rankings in terms of specific items of hypothetical action, the ordinal relations among which are agreed upon by the 18 respondents. It should be emphasized that while the final over-all ranking is fairly consistent with all respondents' rankings, it can not be derived from any one individual's ranking. It is only by the super-imposition of many individual rankings one upon another that the over-all hierarchy may be developed. This does not necessarily mean that no individual in the village would have a conscious awareness of all the differentiations entering into the final hierarchy. Our method, after all, allows the individual to divide all castes into only seven categories (scores from '6' to '0'). However, the question of whether any single informant would grasp the whole hierarchy is worth consideration; perhaps it is only through the addition of the partial perceptions of informants from different *loci* in the hierarchy that the whole picture may be fitted together.

Relations between Responses and Actual Behavior

It is legitimate to ask how everyday behavior corresponds to the ritual purity continuum and the ranking of castes in terms of the continuum. My general impression is that there is a close fit between the ideal patterns (if we may consider the Ritual Distance findings as representing ideal behavior patterns) and actual behavior. Beliefs and behavior have been so adjusted to each other in the Khālāpur subculture, and these patterns are so widely shared by members of all castes, that there is little ambiguity or conflict concerning them.

Household servants such as Water-carriers, Brahman women and Barber women who must touch water vessels, cooking vessels and food of other castes in the course of their duties are not held to be so distant ritually from their clients that they pollute these conductible items. There is little inter-caste sharing of food except at the time of life cycle rites and religious festivals. At these times a Brahman cook is likely to be engaged, from whose hand any other caste may accept either *kachchā* or *pakkā* food or water. At such times and at fairs and bazaars, fried (*pakkā*) food is prepared either by Brahmins or castes of sweetmakers whose ritual distance is not great enough

from most other castes for the food purified by *ghī* or milk to be polluted, when served by them.

The type of food acceptable in payment to servants and artisans has usually been well established between the two castes of employers or clients and servant or artisans. The servant's willingness to take *kachchā* food, *pakkā* food or dry uncooked grain depends upon his perception of his own caste's rank, that of the giver and his own sensitivity to ritual pollution.

The ethnographer observing and participating in village life seldom hears spontaneous discussion of these ritual rules. They are verbalized in the breach, however. Thus, we witnessed an argument once between a Chamār (Untouchable Agricultural Laboring caste) midwife, who had come to serve a Sweeper woman, and the Sweeper relatives concerning her payment. The midwife refused to take even dry grain from persons of a caste she considered inferior to her own, and insisted that her payment be in the form of cow-dung cakes. Our own mistakes sometimes elicited spontaneous assertion of ritual rules for behavior. A Rajput woman once warned us not to touch a basket mounted on a heap of cow-dung cakes in her shed, because the basket belonged to her Sweeper and was consequently contaminating.

An uninformed observer might be misled in his interpretation of what he sees. Thus; a number of women told us of instances in which they allowed some contamination to occur as the lesser of two evils. This might happen when a Sweeper is allowed to come into a high caste woman's courtyard to clean up a baby's urine or feces. The ground where the Sweepess had walked would be spread with cow-dung or sprinkled with purifying water after her departure. Sweepers told us that higher castes did not object to their touching string cots, because at times of marriage Sweepers were pressed into service as porters, carrying the cots from neighborhood houses to the scene of festivities. However, some high caste persons told us that after their cots had been touched in this way, they sprinkled them with Ganges water to purify them again. The Barber woman who was a respondent in the Ritual Distance study told us that she would perform certain services for Chamār Weavers at the time of a marriage, but that she considered herself contaminated by such service and always bathed upon her return home.

Although ritual rules are not verbalized very often in ordinary social intercourse, the rules appeared to be conscious and explicit in the villagers' minds. All interviewees in the Ritual Distance study seemed to be very much aware of the rules and stated general rules by which their caste must abide. They also elaborated upon conditions under which rules might be modified and mentioned changes and relaxation in rules that had occurred in recent years. Most referred to their own behavior in interaction with members of other castes in answering the questions of the Ritual Distance Interview. In other words,

the topic of ritual pollution was clearly meaningful to them and one of which they had conscious control.

The reader should bear in mind, however, that the analysis of the Ritual Distance Interviews was not completed until after my departure from Khālāpur. On the basis of the ideal patterns derived from the questionnaire, a thorough comparison might now be made between ideal and actual behavior.

An obstacle to such a study would be the difficulty of locating and collecting an adequate sample of observations of inter-caste transactions. Even in a period of two seasons in the village, it would have been difficult to observe the four or five thousand interactions about which it was possible to inquire in the interviews (273 questions multiplied by 18 informants).

A further complication lies in the fact that the concern is largely with prohibitions—what should not be done. One can become aware of these most easily through interviewing the participants. Without interpretation from villagers, it would be difficult to know whether behavior in an observed situation is omitted due to ritual prohibitions or due to other requirements or possibilities in the situation. It may not always be appropriate, of course, to ask. Lastly, in observing contact between members of two different castes, it may not always be easy to infer which of the two actors would be polluted if an exchange, say, of an earthenware vessel were to occur between them.

Advantages of the Scaling Method

Some of the possible advantages of this method of studying caste ranking are the following:

1. First is the control which this method makes possible to the investigator over the criteria for ranking, both as stimuli in the interview situation and, later, in the unambiguous meanings of scores and ranks.

2. The scaling technique is a means of systematizing responses from many informants according to an objective test for consensus. Judgment regarding consensus is not dependent upon the impressions of the field investigator or of a few key village informants. And not only is consensus clearly revealed, but lack of consensus is revealed and pinpointed—in this case, both in the hierarchy or actions symbolizing relative status and in the ranking of the castes themselves. On the basis of the perceived differentiations in the ranks of different castes, this method for elucidating evidence from informants may be a means of answering at least in part the question: Are there classes of castes of equal rank or is there a tendency for each caste to have a different rank?

3. If other criteria for ranking are developed into scales, this technique promises to yield some answers to the question of correlation *between* caste rankings according to *different* criteria. If the caste hierarchy in any particular village is a single unitary hierarchy, the scales of items used as status measures for different criteria may "scale together" to form a single scale, and from this single scale a single ranking may be derived.

4. This method can take into account individual or role-based differences in ranking between members of the same caste. Both Srinivas (17) and Guha and Kaul (5) have pointed out sex differences in sensitivity to proximity to other castes. With the scaling method we may locate individual and role-based differences. We need not talk in terms of the customs of a whole caste, e.g., how caste A interacts with caste B, but may speak of the younger women of caste A in contrast to the young men, the older women and the older men. This is true, where the scaling is found to exist in social reality, because individual variations in sensitivity to pollution do not affect the ranking of castes or of items indexing the values upon which ranks are based.

5. Lastly, such control makes objective replication feasible both at different times in the same village and for comparisons between villages in the same region.

SUMMARY

This analysis of the scale-pictures and rankings of castes, derived from the 273 responses to the Ritual Distance Interview by 18 persons, suggests that there are implicit in the cultural ideology of caste uniformities and regularities both in criteria for ritual ranking and in the ritual rankings themselves as shown by the high degree of consensus among these informants. Both the ways in which the ritual hierarchy described here relates to caste rankings based on other criteria and how it relates to the context of daily interpersonal relationships need investigation. However, if the ideal pattern of ritual relationships is well defined, it should be possible to study the deviations from the ideal more effectively.

Manuscript received: April 22, 1959

Revised Manuscript received: October 21, 1959

*Pauline Moller Mahar
Department of Anthropology
University of Arizona
Tucson, Arizona*

APPENDIX

I. The Ritual Distance Interview in English Translation

Now I shall ask you some questions about your customs.

1. Can a touch you? Will you object to it or will he (or they*) become polluted?
2. Can a sit on your cot? Will you object to it . . .
3. Can a come on your cooking area? Will you object to it. . .
4. Can a touch your brass utensils? Will you object to it. . .
5. Can a touch your earthenware vessels? Will you object to it. . .

* The Hindi *wah* can signify either the singular or plural third person in this area.

- † 6. Can a smoke the stem of your pipe? Will you object to it. . . .
 7. Can a smoke the bowl of your pipe? Will you object to it. . . .
 8. Can you eat *pakḥā* (cooked in milk or butter) food from the hand of a?
 9. Can you eat *kachchā* (cooked in water) food such as cooked pulse, bread or rice from the hand of a?
 10. Can you take dry uncooked food (*sukḥā sidhā*) from a?
 † 11. Can you take water from the hand of a?
 12. Can a touch your water vessel? Would you object to it. . . .
 13. Can a touch your children? Would you object to it. . . .

II. Castes of *Khālāpur*—Names in English and Hindi

1. Grainparcher	Bharbhūjā
2. Genealogist	Bhāt
3. Sweeper	Chūrḥā
4. Brahman	Brāhman
5. Chamar (Untouchable agricultural laborer)	Chamār
6. Carpenter	Barhī
7. Washerman	Dhobī
8. Shepherd	Garariyā
9. Gosain (religious sect)	Gusāin
10. Shoemaker	Jatīyā Chamār
11. Water-carrier	Jhīnvar
12. Jogi (religious sect)	Jōgī
13. Kabir Panthi Weaver	Kabirbansi Julāhā
14. Chamar Weaver	Chamār Julāhā
15. Potter	Kumbhār
16. Mirasi (Muslim beggar and entertainer)	Mirāsī Dūm
17. Barber	Nāi
18. Rajput (warrior)	Rājput
19. Muslim Rajput	Musalmān Rājput
20. Goldsmith	Sunār
21. Oilpresser	Tel
22. Merchant	Baniyā

† The wording of items 6 and 11 was changed between the pre-test schedule of items and the survey schedule.

(The name of each caste was filled in the blanks above in turn. All castes were asked about for each question before going on to the next question.)

REFERENCES

1. Blunt, E. A. H., "Commensal and Other Restrictions of Caste," in *The Caste System of Northern India*, Madras: Oxford University Press, 1931.
2. Bogardus, E. S., *Immigration and Race Attitudes*, New York: D. C. Heath, 1928.
3. Dube, S. C., *Indian Village*, Ithaca, New York: Cornell University Press, 1955.
4. Dube, S. C., "Ranking of Castes in Telangana Villages," *Eastern Anthropologist*, 1955, 8, 182-190.
5. Guha, U., and M. N. Kaul, "A Group Distance Study of the Castes of U. P.," *Bulletin of the Department of Anthropology* (Government of India), 1954, 2.
6. Gupta, R., "Caste Ranking and Inter-Caste Relations Among the Muslims of a Village in North Western U. P.," *Eastern Anthropologist*, 1956, 10, 30-42.
7. Guttman, L., "A Basis for Scaling Qualitative Data," *American Sociological Review*, 1944, 9, 139-150.
8. Mandelbaum, D., "The World and World View of the Kota," in McK. Marriott (ed.), *Village India*, The American Anthropological Association, 1955, Vol. 57, No. 3, Part 2, Memoir No. 83.
9. Marriott, McK., "A Technique for the Study of Caste Ranking in South Asia," (unpublished), read at the Fifty-sixth Annual Meeting of the American Anthropological Association, December, 1957.
10. Mayer, A., "Some Hierarchical Aspects of Caste," *Southwestern Journal of Anthropology*, 1956, 12, 117-144.
11. Nadel, S. F., "Caste and Government in Primitive Society," *Journal of the Anthropological Society of Bombay*, 1954, New Series 8, 9-22.
12. Radcliffe-Brown, A. R., "Taboo," in A. R. Radcliffe-Brown, *Structure and Function in Primitive Society*, Glencoe, Illinois: The Free Press, 1952.
13. Riley, M., J. Riley, J. Toby, et al., *Sociological Studies in Scale Analysis*, New Brunswick, New Jersey: Rutgers University Press, 1954.
14. Ryan, B. F., *Caste in Modern Ceylon*, New Brunswick, New Jersey: Rutgers University Press, 1953.
15. Simpson, G. E., and J. M. Yinger, *Racial and Cultural Minorities*, New York: Harper, 1953.
16. Srinivas, M. N., *Religion and Society Among the Coorgs of South India*, London: Oxford University Press, 1952.
17. Srinivas, M. N., "The Social System of a Mystore Village," in McK. Marriott (ed.), *Village India*, The American Anthropological Association, 1955, Vol. 57, No. 3, Part 2, Memoir No. 83.
18. Stevenson, H. N. C., "Status Evaluation in the Hindu Caste System," *Journal Royal Anthropological Institute*, 1954, 4, Part I and II, 45-65.
19. Stouffer, S., L. Guttman, et al., *Measurement and Prediction*, Vol. IV of *Studies in Social Psychology in World War II*, Princeton, N. J.: Princeton University Press, 1950.

Preoccupation with Competitiveness and Social Acceptance Among American and English College Students¹

RALPH H. TURNER, *University of California, Los Angeles*

The principal objective of this paper is to compare the responses of groups of American and English college students on two Guttman type scales. The scales are designed to measure (a) the extent to which the individual sees relations with those about him as competitive and (b) the extent to which he is preoccupied with being liked or socially accepted by those with whom he comes in contact. The secondary objective is to describe the scales, report their characteristics when applied to the American and British cases, and to indicate the background ideas which led to the selection of these two variables for measurement and which determined the particular way in which they were defined.

Attitudes toward competition and popularity are among those personal characteristics in which the English and the American people are often thought to differ widely. Most transparently, they are differently evaluated in the "official" cultures of the two countries. In the United States competition is firmly upheld as an essential virtue by Chambers of Commerce and other civic groups, and a major recent criticism of American public schools has been their insufficient cultivation of competitiveness among the students. The extent of competitiveness in American society became a major preoccupation of elements in the progressive education movement during the inter-war years and inspired such diverse writings as Margaret Mead's (13) study of competition in primitive societies and Slavson's (18) widely read criticism of competitiveness in recreation, and aroused interest in Kropotkin's (10) classic effort to show that mutual aid was as important a principle in evolution as the Darwinian competitive struggle for existence. Recent attempts to describe American "national character," such as those of Mead (12) and Gorer (6, Chs. 3 and 4), have also dwelt upon the pervasiveness of competition in the socializing experiences of the child and in the attitudes of the adult.

¹ U. S. data were collected while the author held a Social Science Research Council Faculty Research Fellowship, with the assistance of a supplementary grant from the Penrose Fund of the American Philosophical Society. British data were collected while the author was a Fulbright Research Fellow at the London School of Economics. Grants from the UCLA Faculty Research Committee paid for statistical processing. The author owes a special debt of gratitude to his colleagues, Wendell Bell and Richard Hill, both of whom read the manuscript painstakingly and offered many valuable suggestions, and to Charles Madge, who arranged access to Birmingham University students.

There appears to be no counterpart in English ideology to American idealization of the competitive way of life. Perhaps the omnipresent queue symbolizes as well as anything the suppression of competition. This feature of British life is so notable that a well-known touring guide (4, p. 487) published in Switzerland advises the foreign motorist who visits Britain: "If people have to wait for anything, they place themselves at the end of a queue and patiently await their turn." In his analysis of British national character Gorer (7, pp. 286 ff.) dwells upon the severe control of any overt expression of aggression as the principle explaining many of his questionnaire findings.

Similarly the preoccupation with popularity, though exalted to the standing of a virtue only since World War I, is often regarded as a strikingly American characteristic. H. S. Commager (3, p. 31) writes of the nineteenth-century American as "usually a good neighbor, not by the English standard of respect for privacy but by the American standard of friendliness." In the same century De Tocqueville (20, I, p. 267) had remarked the lack of "manly candor and independence of opinion" among American leaders, and Francis and Theresa Pulsky (16) had contrasted the touchiness of the American with the proud Englishman who "does not care whether he is liked or disliked by other nations, satisfied that he must be respected." Brogan (1, p. 8) presents the same view of the twentieth century Englishman when he refers to his "indifference to outside criticism," and Gorer (7, p. 51) notes from his recent questionnaire study the "exceptional . . . length of time it takes them to knit new bonds [of friendship]." By contrast, Gorer (6, p. 108 & passim) devotes extended attention to the American preoccupation with popularity, asserting: "The presence, the attention, the admiration of other people thus becomes for Americans a necessary component to their self-esteem, demanded with a feeling of far greater psychological urgency than is usual in other countries." Following the phenomenal sales success of Dale Carnegie's *How to Win Friends and Influence People* (2), David Riesman's (17) depiction of the *other-directed man* has come in many quarters to be synonymous with the typical American.

While the foregoing assertions cannot remain unquestioned in every detail, there seems relatively little basis for doubting the general observation that American culture affords more overt approval to competitive and popularity orientations than does the English culture. But the question of modal individual attitudes in the two countries is not necessarily settled by such cultural generalization. Among the many critics of facile generalizations about national character, Hamilton Fyfe (5) has noted the extent to which individuals vary in every country and remarks especially a tendency to mistake the characteristics of a ruling class for the characteristics of all the people.

In contrast to the differences in overt evaluation of the attitudes under

consideration, there are fundamental similarities in English and American social structures associated especially with the industrial base of the two countries. Just as S. M. Lipset and Reinhard Bendix (11) have suggested that the common industrial base has led to a greater similarity in rates of upward mobility between Britain and America than the surface cultural differences would suggest, so the surface attitudes toward competitiveness and popularity-seeking may have less effect than the underlying structural similarities between the two countries. The less competitive exterior and the surface repudiation of popularity-seeking among the English might conceal private attitudes and preoccupations which are not greatly different from those in the United States.

The preliminary comparison of some American and English students reported in this paper is intended as a small beginning toward assessing the degree to which individual preoccupations correspond with official cultures. American universities reflect the national value in being explicitly more competitive in their organization than their English counterparts. Periodic testing, grading "by the curve," admitting marginal students with the intention of soon eliminating most of them contrast with the prevalent assumption in the English university that a student good enough to be admitted should be able to complete the program satisfactorily and a general freedom from testing until the last year or two of college.² American universities have probably been more resistant to the value of teaching the student how to be popular than have the secondary and elementary schools, but are not immune to the pressures of the environing society.

The two scales were administered first to 1406 university and junior-college students in the United States and their scale characteristics determined from this and a second smaller sample at one of the universities. The scales were then administered to 298 university students in England, principally at Birmingham University, but with a small additional group at Kings College, London. Following a brief discussion of the rationale underlying the scales, we shall present the scale characteristics and compare American and English findings.

THE SCALES AND THEIR RATIONALE

Background in Horney's Treatment of Competition

Although the two variables under consideration seem admirably suited to British-American comparison, the scales were originally devised to serve

² A fuller discussion of differences in competitive emphasis between English and American school systems as related to mobility norms in the two countries will be found in the author's "Sponsored and Contest Mobility," a paper presented at the Fourth World Congress of Sociology, 1959, publication forthcoming in the *American Sociological Review*.

another purpose which will be the subject of a later report. In a general exploration of the impact of the upward mobility experience upon the personality of the mobile individual, attention was turned to the competitive aspects of the experience. The hypothesis emerged that the upwardly mobile person would develop a more pervasive preoccupation with the competitive aspects of his interpersonal relations than would the nonmobile individual. It then seemed reasonable to look to the work of Karen Horney, who had provided a careful statement of the social psychological features of competition while remaining free from the more esoteric assumptions of classical Freudianism. The scale of competitive preoccupation was then designed to correspond in a general way with the configuration that Horney described as "neurotic competitiveness." The relationship suggested by Horney between competitiveness and an exaggerated need for affection inspired the addition of a second scale to deal with the latter. In the light of the impact of Riesman's (17) conception of the other-directed man and of Warner and Abeglen's (21) depiction of the upwardly mobile business executive, there seemed to be important merit in including such a measure. The second scale was also designed to reflect in a general way the configuration described by Horney. While the research has at no time been designed to test Horney's theory, the dependence upon her characterizations of the two variables under investigation requires that we preface the presentation of the scales with a brief résumé of the relevant features of her theory.

Ordinary competitive activity, as defined by Park and Burgess (15, p. 507), in which competition is merely an unsought condition to the pursuit of success is often replaced in modern society by a special kind of attitude which Horney calls neurotic competitiveness. This kind of competitiveness sensitizes the individual to the competitive aspects of all that he does, makes him fearful and hostile toward others as potential competitors, and leads to a vital sense of isolation and loneliness. The loneliness is unbearable, so that the individual tries to escape from his condition primarily through the use of two mechanisms. One of these is the quest for power over others, so that by domination he may lessen the dangers of competitiveness and insure himself against being completely isolated. The other mechanism is the development of an exaggerated need for affection. By seeking to magnify the number and intensity of friendships, the individual strives to insure himself against his loneliness. The pursuit of power and of affection may be effective in the short view, but their very distorted character makes them fail to allay the loneliness from which they arise. In addition, there are inherent contradictions between viewing a man as a competitor and seeking to make a friend of him, as there also are between attempting to dominate him at the same time as one seeks his affection. The greater the extent to which society provokes

competitiveness, the greater the frequency and intensity of such conflicts within the population.

In the present paper we have chosen to dwell upon the competitive attitude and the exaggerated need for affection. No effort has been made to identify the intervening variable of loneliness nor to deal with the quest for power.

The contrast between normal and neurotic competitiveness incorporates three major criteria. The first is a constant tendency to measure oneself against others, irrespective of the character of the situation. The neurotic competitor "measures himself against persons who are in no way potential competitors and who have no goal in common with him" (8, p. 189). Such constant preoccupation transforms competition into an autonomous end, pursued for its own sake apart from any interest in the overt goal at hand.

The neurotic competitor is identified, secondly, by the extent of personal superiority which he seeks. He must be "unique and exceptional," rather than merely better than others. He is likely to experience hero phantasies in exaggerated degree and find it difficult to admit that anyone else might be really as good or as attractive as he.

Third, the neurotic competitor is more preoccupied with seeing his competitor fail than with succeeding himself. What is sometimes a useful technique in competition, namely to undercut the rival, becomes a major concern and an end in itself. Karen Horney describes this tendency as the "implicit hostility in the neurotic's ambitions."

Like competitiveness, the need or search for affection is an attribute of the normal individual, but takes on special qualities in the neurotic. The two characteristics stressed are the "compulsiveness" and the "insatiability" of the demand for affection. Affection is stripped of its ordinary spontaneity and discrimination to become an urgent demand followed indiscriminately. The individual may suffer greatly from failure to win the affection of a stranger or of one among many associates. He may be ready to pay the price of an extreme conformity for the sake of remaining in the good graces of others. On the other hand, the insatiability is shown in a sort of "greediness," a demand for unconditional and exclusive love. Jealousy is easily aroused and there is constant suspicion and testing of the other's affection.

Briefly translated into the terms which fit our present objective, the core relationships may be summarized as follows: In a society in which (a) great value is attached to success both as a symbolic goal and through its instrumental relationship to many other values, (b) the *standards* of success are not well defined by tradition, and (c) rivalry is a major device for motivating individuals in the socialization process, success will, in practice, be judged by *ad hoc* comparisons with others. Consequently, competitive success readily becomes a major preoccupation which colors the individual's experience of his

associates, causing him to resent their successes and constantly to fear for his own relative position. In order to limit the arena of competitive struggle, and because of the socialized dependency on others, the individual anxiously seeks large numbers of friends and intense friendships.

Reconciling Theoretical Contexts

Certain theoretical problems are raised by the use to which we shall put Horney's ideas. First, we shall attempt to draw upon her characterization of neurotic attitudes without concerning ourselves with the distinction between normal and neurotic. Horney's assertion that the neurotic differs from the normal person only in degree, while not followed with complete consistency in her own work, will be our guide in this connection. Second, we shall be translating variables which were originally assumed to lie principally in the unconscious realms into scales measuring only what the respondent can tell about his own attitudes. Although the interpretation of responses according to their accepted and transparent cultural meanings will produce different individual characterizations than will the psychiatrist's reading of hidden meanings into the responses, it is important to note that the purported dynamics of the unconscious depends largely upon analogy from the conscious. Since there is further assumed to be an unbroken continuum from the conscious to the unconscious, there is every reason to anticipate that variables and relationships noted in one realm might be profitably explored in the other.

The broader problem concerns the extent to which selected hypotheses may be imported from a psychoanalytic matrix without commitment to the entire theory. The view we hold is that a major substratum of folk observation and folk generalization which has accumulated through centuries of man's observation of himself underlies the major schools of thought about man and society. These schools have developed not so much from original observations as from selective stress upon certain parts of the fund of folk knowledge. It is in the subsequent elaboration of these starting observations that mutually exclusive explanatory principles are introduced. It is the writer's view that much of Horney's work consists of stripping away the elaborate inventions of the Freudians so as to return more closely to some of the key folk observations from which the system developed. Consequently much of her work is readily reconcilable with other than psychoanalytic contexts.

The Scales

Two Guttman-type scales were devised by preparing a series of statements of varied intensity which would reflect the supposed *continua* from most modest competitiveness and search for affection to the kinds of attitude designated by Karen Horney as neurotic. Each statement was designed simply

to be endorsed or rejected. The statements were then administered to available university classes and revised several times until a fairly adequate pair of eight-item scales for present purposes seemed to have been attained. The items appear in the Appendix.

The first, or *Competitive Preoccupation*, scale consisted of statements ranging from disappointment at doing less well than one's friends to sole preoccupation with trying to beat others competitively. The items were chosen to emphasize competitive attitudes in interpersonal relations, rather than impersonal competitiveness. The scale of *Social Acceptance Preoccupation*, as we have called Horney's need for affection, ranges from concern for close friends to an extreme of distress over the thought that there are people who are not one's friends. Something of the character of the latter scale was indicated by items which were eliminated in the pretests because they failed to scale. A few of these involved a willingness to sacrifice for a friend, suggesting an element of altruism. The rather clear separation from altruistic themes is consistent with the attitude described by Horney.

The principal sample for the study was taken at three schools. At the University of California, Los Angeles, 357 students in a required lower-division history course filled out the questionnaires. Los Angeles City College was selected as a public junior college combining terminal students with students hoping to transfer eventually to a university. The same required history course yielded 747 subjects here. As a private university drawing a more wealthy student body than either UCLA or LACC, Northwestern University was chosen. Questionnaires were administered here to 302 students in an introductory geography class which met a lower-division science requirement. At UCLA and LACC, the samples are fully representative of the lower-division student body, except for absentees. At Northwestern there is probably some underrepresentation of physical science students who will have met their requirements in other than geography courses. Marginal percentages and other scale characteristics have been tabulated separately for the three schools, and the schools have been grouped into a single sample when warranted.

In order to provide a measure of test-retest reliability (or stability) of scores, the questionnaires were administered to the two sections of an introductory physical anthropology course at UCLA and the sixteen items making up the two scales readministered to the same classes a month later. Questionnaires were matched by a system of names and numbers on detachable slips of paper. The 194 questionnaires which could be matched for first and second application made up this sample. The physical anthropology course meets a lower-division requirement in life science and consequently draws a fair cross-section of students, probably undersampling life science students.

The marginal percentages are found in Table 1 for the three administrations just described and for the English sample which will be discussed subsequently. The Competitiveness Scale reveals a satisfactory range of marginal frequencies except for one item with only four per cent endorsement in the main sample. The Acceptance Scale errs in the other direction, containing one item with 95 per cent endorsement.

TABLE 1
Scale Item Characteristics

Item	Percentage of Endorsement				Fourfold Correlation: Test-Retest
	Principal U. S. Sample	2nd U. S. Sample		English Sample	
		Test	Retest		
Competitive Preoccupation Scale					
a	4	2	1	1	— .01
b	12	13	8	15	.54
c	17	13	10	17	.48
d	29	28	25	27	.36
e	41	45	37	42	.28
f	51	47	49	51	.60
g	65	69	78	76	.47
h	83	88	91	89	.49
Social Acceptance Preoccupation Scale					
p	10	7	4	10	.44
q	14	10	15	5	.25
r	19	12	9	16	.49
s	32	27	28	22	.54
t	55	54	56	59	.51
u	73	82	77	78	.42
v	76	77	85	77	.45
w	95	98	99	98	— .01

Some reversals between adjacent items with small differences in the principal sample occur in the reliability samples. The order of items in the principal sample is identical for men and women in both scales. Thus the items on both scales do fall into a uniformly understood order, marred only by certain adjacent pairs in which differences are probably too slight to rule out random reversals.

The distribution of scale scores and crude scores appears in Table 2. Coefficients of reproducibility have been computed on the basis of each of the administrations and are reported in Table 3. The coefficients are uniformly above 90 per cent. Coefficients are practically identical for the two scales based upon the principal sample. The acceptable level of Menzel (14)

TABLE 2

Competitiveness and Acceptance Scores for Principal U. S. and English Samples

Score	Frequency			
	Principal U. S. Sample		English Sample	
	Crude Score	Scale Score	Crude Score	Scale Score
Competitive Preoccupation Scale				
0	69	103	6	9
1	207	196	29	25
2	295	247	65	68
3	292	290	81	69
4	242	234	62	55
5	166	197	35	53
6	85	61	15	10
7	33	46	5	9
8	3	18	0	0
No response	14	14	0	0
Total	1406	1406	298	298
Social Acceptance Preoccupation Scale				
0	16	40	4	5
1	81	100	13	21
2	158	129	39	18
3	336	355	76	73
4	414	431	100	119
5	239	179	41	37
6	103	74	15	13
7	40	45	7	7
8	11	45	3	5
No response	8	8	0	0
Total	1406	1406	298	298

Coefficients of Scalability indicates that the extreme items do not contribute excessively to the reproducibility of the scales. By the criteria of reproducibility and scalability, then, the scales meet usual standards, and are fairly equivalent in their characteristics.

As a further check on the adequacy of the scale characteristics, we have compared the ordering of marginal percentages among the six subsamples (male and female in three schools) which constitute the principal sample. For the competitiveness scale, there are only two instances of reversals between pairs of items, one between items *e* and *f* and one between items *f* and *g*. In the acceptance scale, there are two reversals between *p* and *q*, two between *u* and *v*, and one between *r* and *s*. All reversals are between adjacent

TABLE 3
Scale Characteristics

Characteristic	Principal U. S. Sample	2nd U. S. Sample		English Sample
		Test	Retest	
Competitive Preoccupation Scale				
Number of cases	1392	194	194	298
Number of items	8	8	8	8
Reproducibility	.918	.922	.939	.914
Scalability	.678	.675	.700	.627
Per cent perfect scale types	49	48	59	45
Social Acceptance Preoccupation Scale				
Number of cases	1398	194	194	298
Number of items	8	8	8	8
Reproducibility	.919	.942	.950	.935
Scalability	.630	.677	.713	.513
Per cent perfect scale types	51	61	68	58

items with small differences in marginal frequency. Both scales—and the acceptance scale especially—err through the inclusion of some items whose marginals are too nearly the same.

Reliability of the sort which is tested by the split-half method is ordinarily presumed to have been measured through the reproducibility coefficient, but the same cannot be said of retest reliability. Retest reliability is a crucial criterion for any scale which is supposed to measure relatively stable attributes of the individual, such as native intelligence. The two variables under measurement here are of a sort assumed to be less invariant—more responsive to the situation and to events of the recent past. At the same time, however, they are assumed to be well entwined in the total organization of personality and consequently unlikely to change radically within a short space of time. Thus the scales are of dubious merit unless they show a substantial degree of retest stability, but the standards of reliability are not so high as for measures of capacities and aptitudes or even for major values.

Retest stability was measured by two applications of the scales to an introductory anthropology class, one month apart. One month was felt to be sufficient time to minimize the memory of answers to specific items, though it was anticipated that the lack of "freshness" on second application might have some consistent effects on the responses. Administrations were planned to avoid weeks during which examinations were usually being given because of the possible effect of such a contingency on competitiveness scores. In order further to reduce the possible impact of memory and also to eliminate any special arrangement effect, the order of all the items was changed on the

second administration. In the first administration, as in the principal U. S. sample and the English sample, the acceptance scale had been presented first. On retest the competitiveness items were presented first. In the first administration, the same random order of items within the scales was used as in the principal U. S. sample and in the English sample. On retest the order of items was again randomized with respect to the first application. Thus by adding to the month interval a change in order of both scales and items, the scales were subjected to the most stringent kind of test of stability.

The product-moment correlation coefficients between test and retest were .59 and .65 for the competitiveness and acceptance scales, respectively.³ Phi coefficients were also computed for each of the 16 items and have been reported in Table 1. The latter cluster near to .50, except in case of one extreme item in each scale and three others which are somewhat lower. A considerable degree of stability is indicated by these coefficients, but not enough to suggest that an inflexible attribute of personality is involved. The coefficients fall within the range that had been anticipated from our discussion of the nature of the variables under examination.

There remains the possibility that some consistent practice effect might show itself. Accordingly the mean test and retest scores have been compared for each scale. The differences are very small and far short of even a ten per cent confidence level. An examination of marginals does, however, suggest a practice effect on the competitiveness scale. Items which received high endorsement on first application received even higher endorsement the second time and those which received low rates of endorsement were endorsed even less frequently the second time. Thus, while there is no practice effect on average over-all scores, there may be a polarizing tendency with respect to items on retest.

In the absence of any independent test of validity, it is useful in clarifying the nature of the scales to note their relationship to two questions not otherwise incorporated into the present report. The term "preoccupation" has been employed to designate the scales, since they are not supposed to measure the extent to which success or having friends is important to the subjects, but the extent to which these concerns suffuse their relations with others. We have already noted that altruistic items clearly did not scale on the same dimension as our social acceptance preoccupation. Similarly, a measure of the desire to excel, which we shall call *eminence aspiration*, is provided by a question included in the smaller United States sample. Subjects were asked, "After

³ Although scale scores have been computed and punched on cards for all subjects, crude scores have been used throughout the paper in correlational analysis and mean comparisons. The choice was made because of a preference for interpreting imperfect scale types as departures from unidimensionality rather than as respondents' errors.

you are in the occupation which will be your life work, when will you consider yourself successful enough that you can relax and stop trying so hard to get ahead?" The seven responses among which they were to choose ranged from "1. When I am doing well enough to stay in the occupation," to "6. When doing better than everyone else in my occupation," to "7. Never."⁴ There is an absence of correlation or a very minor negative correlation between *eminence aspiration*, thus measured, and both of the scales. Thus the drive to excel in one's chosen field, as one form of ambition, is something different from a preoccupation with the specifically competitive aspects of one's relations with others.

Since both of the scales were derived from descriptions of neurotic attitudes, they should contain considerable overtones of anxiety. One question asked of the principal United States sample was the following: "Do you ever have any doubts about whether you will achieve all of the success you are aiming for?" The five alternatives ranged from "I never have a moment of doubt," to "I am always in doubt." The responses yielded small correlations (from .11 to .22) with both of the scales. Thus there appears to be a significant but very slight tendency for lack of confidence in achieving "success" to be associated with high scores on both of the scales.

In summary, the two scales stand up fairly well to the usual tests, showing some defects but generally functioning in such a fashion as to warrant their further use until improved scales can be devised. The major criterion not explored is that of validity. No effort was made to test validity apart from the assumption that items measured what they seem to measure when taken at face value. All inferences from the scales must rest upon the untested assumption of validity. We can say only that both scales probably are unrelated to striving to excel in one's chosen field and are probably slightly related in a negative direction to the confidence one has in achieving desired success.

FINDINGS FOR ENGLISH AND AMERICAN STUDENTS

Variation and Correlates among American Students

In order to gain a further idea of the characteristics of the scales and to provide a basis for more adequate comparison between national groups, we

⁴ The *eminence* question regularly yields a bimodal distribution, with response numbers 4 and 5 ("When doing much better than the average in my occupation," and "When recognized as one of the top persons in my occupation") and "Never" being most frequently chosen. Men scored very significantly higher than did women when the latter answered a companion question which asked when they thought their "future husband should relax and stop trying so hard to get ahead." In light of the subsequent portion of this paper, it is of interest to note that the English students did not differ appreciably from their UCLA counterparts in their response to this question, and that the same difference between the sexes was found in both national groups.

shall first record the intercorrelation between scales and the variations by sex and school.

Pearson product-moment correlations between the two scales have been computed for males alone, females alone, and the two combined. The resulting coefficients are .28, .37, and .27, respectively, and are all significantly greater than zero beyond the one per cent level of confidence. The positive correlations are to be expected according to the theory of Karen Horney, but the small size of correlations makes it impossible to rule out the possibility that these are spurious reflections of acquiescence bias.

Although the scales are correlated, we can demonstrate that the respondents do discriminate between them in their answers. In the smaller United States sample the intercorrelation of .29 between scales on first application can be compared with the correlations of .59 and .65 between test and retest for the competitiveness scale and the acceptance scale, respectively. Each of the latter coefficients is significantly greater than the intercorrelation between scales at well beyond the one per cent confidence level. Thus students are significantly more likely to secure the same scores when answering the same scale at a month interval than when answering different scales at the same sitting.

Variations by sex and school are reported in Table 4. For each of the three schools the mean competitiveness score for men is significantly higher than the mean score for women (in each case $P < .01$). For the acceptance scale there are no significant differences. The higher rate of competitiveness preoccupation for men is in keeping with what might have been expected in light of the more active involvement of men than of women in occupational and athletic competition. However, the rather unmasculine evaluation of overconcern for the good will of others might have led us to expect that women would score higher on the second scale than men, except for the compensating

TABLE 4

*Mean Competitiveness and Acceptance Scores for Principal U. S. Sample **

School	Male	Female
Competitive Preoccupation Scale		
LACC	3.17	2.62
UCLA	3.57	2.57
Northwestern	3.53	2.73
Social Acceptance Preoccupation Scale		
LACC	3.87	3.77
UCLA	3.64	3.78
Northwestern	3.57	3.61

* All scores are crude scores, i.e., the number of items endorsed by the subject.

effect of the intercorrelation between the scales. To ascertain whether there is any such hidden tendency for women to be more preoccupied with social acceptance than men requires the use of partial correlation. If the observed relationship between competitiveness and sex be defined as positive, then it would be essential to demonstrate a negative relation between sex and acceptance preoccupation when competitiveness is held constant. Since sex is a dichotomy, biserial and product-moment zero-order coefficients must be combined to compute the desired coefficient of partial correlation. Some statistical error is introduced by this expedient, so the result can be regarded only as a broad approximation. With biserial coefficients of .017 between acceptance preoccupation and sex and .250 between competitiveness and sex, the partial coefficient between the acceptance scale and sex is $-.055$. Since the coefficient approximates zero, we must conclude that no hidden relationship between social acceptance preoccupation and sex has been located.

A comparison of scores among schools shows no difference in competitiveness for females and no difference in acceptance preoccupation for either sex. UCLA and Northwestern males do not differ significantly in competitive preoccupation, but both are significantly more competitive than are LACC males ($P < .01$ and $P < .05$, respectively). There is ready justification for either of two opposing hypotheses regarding male competitiveness and the schools. The view that competitive orientation is enhanced by the mobility experience would lead to the hypothesis that LACC men with their lower family status backgrounds should be most competitive. On the other hand the view that competitiveness is directly related to level of ambition would have led to a prediction consistent with the observed findings. Other characteristics of the schools may have accounted for the finding, and preliminary exploration fails to show a simple relationship between background or ambition and competitiveness among the individual subjects.

Findings for English Students

The opportunity to administer the same questionnaire to some groups of English university students came the year after the data just reviewed had been collected. Although conditions precluded securing adequate or even equivalent samples, it seemed worth while to look for any gross differences between the English students who were available and the American students, and to determine whether the characteristics of the scales followed similar lines to those in the United States. Questionnaires were administered by the investigator to the large departmental seminars in commerce (and social science), in engineering and in education at the University of Birmingham, and to a smaller group of students at Kings College, London. Except for the first group, these were seminars required of all students in the curriculum.

The commerce seminar was directed to all students in the curriculum, but attendance was not compulsory. Birmingham is more nearly equivalent to UCLA than to other schools. It is one of the best established of the universities outside of the old prestige circuit.

The English data will be examined in the following steps: First, we shall determine whether the instruments retain their scale qualities when applied to an English sample. Second, the interrelation of the two scales will be examined to determine whether they bear the same positive intercorrelation as in the United States sample. Third, British and American scores will be compared to ascertain whether the culturally indicated differences are present.

The listing in Table 1 reveals a similar ordering of items according to marginal percentages in the English and American samples. There are no reversals of items on the competitive scale. The two reversals on the acceptance scale are between adjacent items which differ very little and for which reversals had occurred in one of the American samples. In addition, the percentages themselves are strikingly close to those found in the American samples. A further examination of percentages when tabulations were made separately by curriculum and sex showed only a few reversals of adjacent items with small differences in marginal frequency. Reproducibility coefficients (Table 3) are likewise similar to those secured from American samples. The coefficient of scalability for the competitiveness scale is satisfactory, but for the social acceptance scale it falls below the suggested minimum. The conclusion seems clear that the competitiveness scale, developed and tested with American college students, is no less adequate in its scaling characteristics for the English students sampled. The social acceptance scale, on the other hand, is somewhat less adequate as a device for ordering the English students, when appropriate correction is made for extreme items. The English students' preoccupations may be less sharply organized about a dimension of gaining and keeping social acceptance, but considerable ordering is still indicated. The anticipation that the items might be meaningless or that most items would be so universally rejected as to destroy the discrimination among items was unfounded.

Product-moment correlations between the scales are .17 for English males, .33 for English females, and .20 for the sexes combined. Although the coefficients are consistently a little lower than those for the principal United States sample, differences this great would frequently occur by chance. The relationship is also greater for women than for men, as it was in the American sample, though again the differences are not large enough to minimize chance as an explanation.

The remaining comparisons can be drawn from the mean scores presented in Table 5. Means are reported separately for the subsamples because

TABLE 5
*Mean Competitiveness and Acceptance Scores for English Students**

Curriculum	Male	Female
Competitive Preoccupation Scale		
Engineering	3.39	—
Teacher training	3.28	2.65
Commerce and Social Science	3.43	2.86
Social Acceptance Preoccupation Scale		
Engineering	3.90	—
Teacher training	3.69	3.70
Commerce and Social Science	3.39	3.67

* All scores are crude scores, i.e., the number of items endorsed by the subject.

together they do not approximate a representative sample of the University population. Because they do represent contrasting groups of students, however, they should reveal any important contrasts with American students which would apply to a more adequate sample. First, we note that the pattern of sex differences is identical to that in the United States. There is no difference in acceptance scores, but there are differences in both of the curricula for which both men and women students were included ($P < .05$ in each instance). Second, there are no significant differences between students in the three curricula with respect to either scale, except that engineering students may be a little more preoccupied with social acceptance than commerce and social science males ($P < .025$).⁵ Finally, it is to be noted that there are no significant differences between English and American subsamples on either scale.

The upshot of the international comparison is to note that, in most important respects, the English and the American students might have been samples randomly drawn from a common universe. The scales have similar applicability, the distribution of attitudes is the same, the differences between sexes are the same, and the interrelation of attitudes is similar. Within the limitations of the methods used, the results one-sidedly favor the view that surface cultural differences in competitive attitude and in striving to be liked by everyone do not fundamentally alter the private preoccupation with these matters induced by a similar social and economic structure.

⁵ A difference of such doubtful significance would not merit notice except that there may be a parallel with the American data. Engineering is relatively poorly established in British universities and carries low status, the word "engineer" more often referring to a type of skilled craftsman than to a professional in common usage. Engineering students are drawn heavily from working class backgrounds. LACC is similarly a relatively low status collegiate institution drawing heavily from working class backgrounds and reveals the highest mean acceptance preoccupation for males.

SUMMARY

Karen Horney, in her discussion of contemporary neurosis, has called attention to a tendency toward heightened preoccupation with *competitiveness* as a frequent aspect of modern society. She has further suggested that one way in which people seek to escape the unpleasant consequences of universal competitiveness is through an exaggerated search for love. Commentaries on American and English cultures frequently cite these two attitudes as being much more prevalent in the United States. However, it is possible that the underlying similarities of social structure in the two countries may generate very similar preoccupations which are concealed but not eradicated. Two eight-item Guttman-type scales have been devised to measure manifest preoccupation with competitiveness and with social acceptance (love-seeking), and have been administered to samples of American and English college students. The scales show acceptable reproducibility in both national samples, though the acceptance scale is deficient in the English sample by the scalability test. In the United States, both scales show moderate stability under stringent retest conditions. The two scales reveal low but significant positive intercorrelations in both national samples. The competitiveness scores (but not the social acceptance scores) are higher for men than for women in both national groups. The American four-year university men show higher competitiveness scores than those from a public junior college. In general, the results suggest that the scales are useful, and that English and American students are similar with respect to these preoccupations in spite of surface cultural differences.

Manuscript received: August 24, 1959

Revised Manuscript received: November 2, 1959

Ralph H. Turner

Department of Anthropology and Sociology

University of California

Los Angeles 24, California

REFERENCES

1. Brogan, D. W., *The English People: Impressions and Observations*, New York: Alfred A. Knopf, 1943.
2. Carnegie, D., *How to Win Friends and Influence People*, Philadelphia: Universal Book and Bible House, 1936.
3. Commager, H. S., *The American Mind: An Interpretation of American Thought and Character Since the 1880's*, New Haven: Yale University Press, 1950.
4. *Europa Touring: Motoring Guide of Europe*, Berne, Switzerland: Hallwag Berne, 1957.
5. Fyfe, H., *The Illusion of National Character*, London: Watts and Co., 1946.
6. Gorer, G., *The American People: A Study in National Character*, New York: W. W. Norton and Co., 1948.

7. ———, *Exploring English Character*, London: The Cresset Press, 1955.
8. Horney, K., *The Neurotic Personality of Our Time*, New York: W. W. Norton and Co., 1937.
9. ———, *New Ways in Psychoanalysis*, New York: W. W. Norton and Co., 1939.
10. Kropotkin, P., *Mutual Aid: A Factor of Evolution*, New York: McClure Phillips and Co., 1902.
11. Lipset, S. M., and R. Bendix, *Social Mobility in Industrial Society*, Berkeley and Los Angeles: University of California Press, 1958.
12. Mead, M., *And Keep Your Powder Dry*, New York: Morrow, 1942.
13. Mead, M., ed., *Cooperation and Competition Among Primitive Peoples*, New York: McGraw-Hill Book Co., 1937.
14. Menzel, H., "A New Coefficient for Scalogram Analysis," *Public Opinion Quarterly*, 1953, 17, 268-280.
15. Park, R. E., and E. W. Burgess, *Introduction to the Science of Sociology*, Chicago: University of Chicago Press, 1921.
16. Pulszky, F., and T. Pulszky, "The British and American Characters Compared," in H. S. Commager, ed., *America in Perspective*, New York: Mentor Books, 1947, pp. 109-114.
17. Riesman, D., et al., *The Lonely Crowd*, New Haven: Yale University Press, 1950.
18. Slavson, S. R., *Recreation and the Total Personality*, New York: Association Press, 1948.
19. Stouffer, S. A., et al., *Measurement and Prediction*, Princeton: Princeton University Press, 1950.
20. de Tocqueville, A., *Democracy in America*, New York: Alfred A. Knopf, 1945.
21. Warner, W. L., and J. C. Abegglen, *Big Business Leaders in America*, New York: Harper and Brothers, 1955.
22. Williams, R., *American Society: A Sociological Interpretation*, New York: Alfred A. Knopf, 1951.

APPENDIX

Wording of Scale Items

Competitive Preoccupation Scale

- a. The only thing I think of all day long is how to outdo somebody in some way.
- b. It often spoils my sense of accomplishment to find that a few of my friends are doing as well as I am.
- c. I usually feel a little uncomfortable with an acquaintance who I know can outdo me in one or two things.
- d. I am always somewhat disappointed when anything I do isn't distinctly better than the great majority of my friends and associates.
- e. I am always somewhat disappointed when anything I do isn't distinctly better than the average of my friends and associates.
- f. A little zest is always added to any accomplishment if I am also outdoing some friend or associate.
- g. I often find that I am unintentionally comparing myself with others.
- h. I am always at least a little disappointed when anything I do isn't quite as good as the average of my friends and associates.

Social Acceptance Preoccupation Scale

- p. I can't stand the thought that there are people who aren't my friends.
- q. The only thing that really gives me any satisfaction is knowing that I have made someone take a liking to me.
- r. The sense of accomplishment and interest in any task I undertake is never so important to me as the thought that doing the job well might get somebody to like me better.
- s. I can't enjoy any accomplishment at all if I make someone dislike me in the process.
- t. I often find that I am unintentionally thinking about whether my actions are getting people to like me or to dislike me.
- u. I am always at least a little disappointed when I meet someone who doesn't seem to like me.
- v. I am likely to feel at least a little disappointed if an acquaintance of mine turns against me.
- w. I am likely to feel at least a little disappointed if a good friend of mine turns against me.



